# Micro Theory

Geoffrey Heal

August 22, 2017

## Contents

I  **Choice under constraint**  
1  **Consumer Choice**  
   1.1  Comparative Statics  
   1.2  Preferences  
2  **Utility Maximization Problem (UMP)**  
   2.1  Digression on constrained maximization  
   2.2  UMP solutions  
3  **Expenditure Minimization Problem (EMP)**  
   3.1  Using First Order Conditions  
   3.2  Derivation of Slutsky Equation for labor supply  
      3.2.1  Example  
   3.3  Welfare Evaluation of Economic Changes  
   3.4  Deadweight loss from commodity taxation.  
4  **First Problem Set**  
5  **Preference Aggregation**  
6  **Preference aggregation and social choice**
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Final comments on preferences and choice</td>
<td>34</td>
</tr>
<tr>
<td>7.1</td>
<td>Framing</td>
<td>34</td>
</tr>
<tr>
<td>7.2</td>
<td>Endogenous Preferences</td>
<td>35</td>
</tr>
<tr>
<td>II</td>
<td>Firms and Production Plans</td>
<td>36</td>
</tr>
<tr>
<td>7.3</td>
<td>Properties of the Production Set</td>
<td>37</td>
</tr>
<tr>
<td>7.4</td>
<td>Profit Maximization (PMP)</td>
<td>38</td>
</tr>
<tr>
<td>7.5</td>
<td>Cost Minimization (CMP)</td>
<td>39</td>
</tr>
<tr>
<td>7.5.1</td>
<td>Examples:</td>
<td>40</td>
</tr>
<tr>
<td>7.6</td>
<td>Aggregation of Firms</td>
<td>44</td>
</tr>
<tr>
<td>8</td>
<td>Second Problem Set</td>
<td>45</td>
</tr>
<tr>
<td>III</td>
<td>Choice under Uncertainty</td>
<td>46</td>
</tr>
<tr>
<td>9</td>
<td>Preferences over Lotteries</td>
<td>46</td>
</tr>
<tr>
<td>9.1</td>
<td>Paradoxes</td>
<td>50</td>
</tr>
<tr>
<td>10</td>
<td>Risk Aversion</td>
<td>52</td>
</tr>
<tr>
<td>10.1</td>
<td>Risk Management</td>
<td>53</td>
</tr>
<tr>
<td>10.1.1</td>
<td>Mean-Variance</td>
<td>58</td>
</tr>
<tr>
<td>10.2</td>
<td>Comparison of payoffs in terms of return and risk</td>
<td>59</td>
</tr>
<tr>
<td>10.3</td>
<td>A Geometric Approach to Insurance</td>
<td>61</td>
</tr>
<tr>
<td>11</td>
<td>Discount Rates and the Elasticity of Marginal Utility</td>
<td>62</td>
</tr>
<tr>
<td>12</td>
<td>State-Dependent Preferences</td>
<td>63</td>
</tr>
<tr>
<td>13</td>
<td>Subjective Probabilities</td>
<td>65</td>
</tr>
<tr>
<td>13.1</td>
<td>Savage’s Axioms</td>
<td>66</td>
</tr>
<tr>
<td>14</td>
<td>Non-Expected Utility Approaches</td>
<td>68</td>
</tr>
<tr>
<td>14.1</td>
<td>MinMax Approaches</td>
<td>69</td>
</tr>
<tr>
<td>14.2</td>
<td>MaxMin Expected Utility</td>
<td>70</td>
</tr>
<tr>
<td>14.3</td>
<td>Smooth Ambiguity Aversion</td>
<td>71</td>
</tr>
</tbody>
</table>
14.4 Examples ........................................... 71

15 Third Problem Set ................................. 77

IV General Equilibrium ......................... 78

16 Edgeworth Box ................................ 79

17 The Theorems of Welfare Economics ...... 82

18 Problem Set 4 .................................... 87

19 Existence of General Equilibrium .......... 88

20 Public Goods .................................. 89

   20.1 A Simple Case ............................... 90

   20.1.1 Digression on Game Theory .......... 91

   20.2 Private Provision in the Simple Case .. 92

   20.3 The General Case ......................... 95

21 External Effects .............................. 99

22 Common Property Resources ............... 100

23 Non-Convex Production Sets ............... 102

24 Time and Uncertainty in General Equilibrium: the Arrow-Debreu Model ............. 104
These notes are loosely based on *Microeconomic Theory* by Andreu Mas-Colell, Andrew Whinston and Jerry Green. But the coverage differs: I cover some material that they don’t and they cover a lot that I don’t. In class presentations I will supplement these notes with a lot of diagrams and intuitive explanations.

There are three basic topics to be covered this semester:

1. Optimal constrained choice under certainty - consumers, firms, governments, investors

2. Optimal constrained choice under uncertainty and incomplete information - applying this to all of the above topics again

3. General market equilibrium: interaction of consumers, firms, governments, investors via markets
Part I
Choice under constraint

Examples are

1. Consumer maximizing wellbeing subject to budget or wealth constraints
2. Firm maximizing profits subject to technological and market constraints
3. Government maximizing welfare subject to budget, technology constraints

Basic concepts are choice set and preference ordering. A consumer wants to make the best choice from the set of alternative available to her (choice set). This requires that we define the set of alternatives to be considered and also a way of ranking these (preference ordering) so that we can define a “best choice.”

Definition 1. The Choice Set is the set of alternatives from which consumer can choose. This is defined by what is offered and technological constraints. Denote this by $X$.

Next we need

Definition 2. Preference Ordering: a ranking of all the alternatives in $X$. Denoted by $\succeq$, $x \succeq y$ denotes “$x$ is preferred or indifferent to $y$.”

Strict preference defined by $x \succeq y$ and not $y \succeq x$, denoted $x \succ y$.

Indifference defined by $x \succeq y$ and $y \succeq x$, denoted $x \sim y$.

Complete means for all $x, y \in X$, either $x \succeq y$ or $y \succeq x$.

Transitive means $x \succeq y, y \succeq z \Rightarrow x \succeq z$.

A preference ordering is said to be rational if it is complete and transitive.

Proposition 1. If $\succeq$ is rational then

(1) $\succ$ is transitive and irreflexive ($x \succ x$ is never true)
(2) $\sim$ is reflexive ($x \sim x \forall x$) and transitive and
(3) if $x \succ y \succeq z$ then $x \succ z$. 


Proof. First prove that $\succ$ is transitive (by contradiction). $x \succ y, y \succ z \Rightarrow x \gtrsim y, y \gtrsim z \Rightarrow x \gtrsim z$. Suppose it is not the case that $x \succ z$. Then $x \sim z$. This means that $x \gtrsim z$ and $z \gtrsim x$. But $z \gtrsim x, x \gtrsim y \Rightarrow z \gtrsim y$. Contradicts $y \succ z$.

To show that $\succ$ is irreflexive note that $x \succ y \Rightarrow x \gtrsim y$ & not $y \gtrsim x$ so $x \succ x \Rightarrow x \gtrsim x$ & not $x \gtrsim x$ a contradiction.

Next prove that $\sim$ is transitive. Assume $x \sim y, y \sim z$. Then $x \gtrsim y, y \gtrsim x \Rightarrow z \gtrsim x$. And $x \gtrsim y, y \gtrsim z \Rightarrow x \gtrsim z$. So $z \gtrsim x$ and $x \gtrsim z \Rightarrow z \sim x$.

Finally consider the last case, $x \succ y \gtrsim z$. There are two possibilities, $y \succ z$ and $y \sim z$. In the first case we are done by the transitivity of strict preference. So assume that $x \succ y \sim z$.

Also assume that $x \lesssim z$, meaning that either $x \sim z$ or $x \prec z$. But if $x \sim z$ and $z \sim y$ then $x \sim y$ which contradicts $x \succ y$. And if $x \prec z$ then $z \gtrsim x$ and also $z \lesssim y$ (because $z \sim y$) so $y \gtrsim x$ a contradiction. \[\square\]

1 Consumer Choice

Commodity space is $\mathbb{R}^N$. Commodity bundles are vectors in $\mathbb{R}^N$. Consumption set $X \subseteq \mathbb{R}^N$. Reflects physiological constraints - how much effort can be supplied, how much food is needed to survive, dependence of effort on food consumption, maximum of 24 hours in a day, etc. Generally closed, for simplicity take to be $\mathbb{R}^N_+$ and to be the same as the choice set.

Budget is constraint on amount that can be spent, not to exceed income $W \in \mathbb{R}$. $p \in \mathbb{R}^N$ is a price vector, and $p.x$ is the cost of bundle $x$ at prices $p$. The budget set $B$ is $B = \{x \in \mathbb{R}^N : p.x \leq W\}$.

Consumer problem is to choose a best point in $B$ according to the ordering $\gtrsim$, i.e. to choose in $\{x \in B : \forall y \in B, x \gtrsim y\}$. The demand correspondence is $x(p,W) = \{x \in \mathbb{R}^N_+ : x \in B, \forall y \in B, x \gtrsim y\}$. This is a correspondence (set-valued map) as it is not necessarily single-valued.

Note that the price vector $p$ is orthogonal to the budget hyperplane $H = \{x : p.x = W\}$, i.e. $p.z = 0$ for any vector $z$ lying in the hyperplane $H$.

To see this let $y$ be such that $p.y = W$, and also let $x \in H$, then $(x - y)$ is a vector in $p.x = W$, and $p.x = p.y = W$ so $p.[x - y] = 0$ and $p$ is orthogonal to $(x - y)$.

Definition 3. Demand correspondence $x(p,W)$ is homogeneous of degree zero if $x(p,W) = x(\alpha p, \alpha W)$ for any $p, W$ and any $\alpha > 0$. 

6
Also needed is

\textbf{Definition 4.} Demand correspondence satisfies Walras’ Law if for every \( p \geq 0 \) and \( W > 0 \) we have \( p.x = W \) for all \( x \in x(p,W) \).

Assume that the demand correspondence is single-valued, i.e. a function.

\textbf{Proposition 2.} The demand function is homogeneous of degree zero.

\textit{Proof.} \( \{ x : p.x \leq W \} = \{ x : \alpha p.x \leq \alpha W \} \)

\subsection{1.1 Comparative Statics}

We investigate how consumption changes with changes in wealth and in prices. For fixed \( \bar{p} \), the function \( x(\bar{p},W) \) is the\textit{ Engel function}. Its image in \( \mathbb{R}^n \), \( \{ x(\bar{p},W) : W > 0 \} \) is the\textit{ wealth expansion path}. A commodity \( l \) is normal at \( (p,W) \) if \( \partial x_l/\partial W \geq 0 \); otherwise it is inferior.

Locus of \( x_i(p,W) \) as \( p_i \) varies, all other variables constant, is the\textit{ demand curve} for good \( i \), sometimes known as the offer curve.

The derivative \( \partial x_i(p,W)/\partial p_i \) is the\textit{ own price effect}, and the cross derivatives \( \partial x_i(p,W)/\partial p_k \) are the\textit{ cross price effects}. There is a matrix of such cross partials with the own effects on the diagonal

\[
D_p(p,Y) = \begin{bmatrix}
\frac{\partial x_1(p,W)}{\partial p_1} & \cdots & \frac{\partial x_1(p,W)}{\partial p_N} \\
\vdots & \ddots & \vdots \\
\frac{\partial x_N(p,W)}{\partial p_1} & \cdots & \frac{\partial x_N(p,W)}{\partial p_N}
\end{bmatrix}
\]

Elasticities of demand are defined as

\[
\varepsilon_{n,k} = \frac{\partial x_n(p,W)}{\partial p_k} \frac{p_k}{x_n(p,W)} \\
\varepsilon_{n,W} = \frac{\partial x_n(p,W)}{\partial W} \frac{W}{x_n(p,W)}
\]

\textbf{Proposition 3.} If the demand function \( x(p,W) \) is homogeneous of degree zero then for all \( p \) and \( W \)

\[
\sum_{k=1}^N \frac{\partial x_n(p,W)}{\partial p_k} p_k + \frac{\partial x_n(p,W)}{\partial W} W = 0 \quad \forall n = 1,\ldots,N
\]

\footnote{This means \( p_i \geq 0 \forall i \) and \( p_i > 0 \) some \( i \).}
Proof. Consider the demand function $x(\alpha p, \alpha W)$, differentiate with respect to $\alpha$ and let $\alpha = 1$. 
\[ \frac{\partial x}{\partial \alpha} = \sum k \frac{\partial x}{\partial p_k} \frac{\partial p_k}{\partial \alpha} + \frac{\partial x}{\partial W} \frac{\partial W}{\partial \alpha} = \left( \frac{\partial x}{\partial p_1}, \ldots, \frac{\partial x}{\partial p_N} \right) (p_1, \ldots, p_N) + \frac{\partial x}{\partial W}. \]
But as $x(p, W) = x(\alpha p, \alpha W) \forall \alpha > 0$, $\frac{\partial x}{\partial \alpha} = 0$.

In matrix notation $D_p x(p, W) p + D_W x(p, W) W = 0$.

By dividing through by $x_n$ we see that in elasticity form proposition 3 takes the form
\[ \sum_k \varepsilon_{n,k}(p, W) + \varepsilon_{n,W}(p, W) = 0 \forall n \]
so the sum of all elasticities pof demand for a good is zero.

**Proposition 4.** From Walras’ Law - $p.x(p, W) = W$ - it follows that $\sum_n p_n \frac{\partial x_n}{\partial p_k} + x_k = 0 \forall k$ and $\sum_n p_n \frac{\partial x_n}{\partial W} = 1$.

**Proof.** Just differentiate Walras’ Law by prices or by wealth. \( \square \)

### 1.2 Preferences

We have assumed completeness and transitivity of preference orderings. We now add assumptions about desirability and convexity.

**Definition 5.** The preference ordering $\succeq$ on $X$ is **monotone** if $x \in X \& y \gg x$ implies $y \succ x$.\(^2\) So if $y$ has more of every good than $x$, $y$ is better. It is strongly monotone if $y \succeq x \& y \not= x$ implies $y \succ x$. In this case $y$ only has to have more of some good.

A weaker version is

**Definition 6.** The preference relation $\succeq$ on $X$ is **locally non-satiated** if for every $x \in X$ and every $\epsilon > 0$ there is a $y \in X$ such that $\|x - y\| < \epsilon$ and $y \succ x$.\(^3\)

The **indifference set** containing $x$ is $\{ y \in X : y \sim x \}$. 

The **upper contour set** is the set of points at least as good as $x$, that is the set of points on or above the indifference set containing $x$, $\{ y : y \succeq x \}$. 

The **lower contour set** is the set of points that $x$ is at least as good as: $\{ y : x \succeq y \}$. 

\(^2\) $x \gg y$ means that $x_i > y_i \forall i$.

\(^3\) $\| x - y \| = \sum_i (x_i - y_i)^2$
Local non-satiation rules out thick indifference sets, as do the various forms of monotonicity.

A set \( S \) is \textbf{convex} if whenever \( x, y \in S \), \( \lambda x + (1 - \lambda) y \in S \forall \lambda \in [0, 1] \). It is strictly convex if in addition \( \lambda x + (1 - \lambda y) \) is not in the boundary of \( S \).

\textbf{Definition 7.} The preference relation \( \succeq \) on \( X \) is convex if for every \( x \in X \) the upper contour set is convex, that is if \( y \succeq x \& z \succeq x \) then \( \alpha y + (1 - \alpha) z \succeq x \) for any \( \alpha \in [0, 1] \).

A strengthening of this is

\textbf{Definition 8.} The preference relation \( \succeq \) on \( X \) is \textbf{strictly convex} if for every \( x \in X \) the upper contour set is strictly convex or if we have if \( y \succeq x \& z \succeq x \& y \neq z \) then \( \alpha y + (1 - \alpha) z \succ x \) for any \( \alpha \in (0, 1) \).

Homothetic preferences are often used because they have some neat properties:

\textbf{Definition 9.} A monotone preference relation on \( X \) is \textbf{homothetic} if \( x \sim y \Rightarrow \alpha x \sim \alpha y \forall \alpha \geq 0 \)

Intuitively, with homothetic preferences every indifference curve can be obtained from one by scaling it up or down. Another property often used is that of being quasi-linear:

\textbf{Definition 10.} A preference relation \( \succeq \) is \textbf{quasi-linear} with respect to commodity 1 (called the numeraire commodity) if

\begin{enumerate}
  \item all indifference sets are parallel displacements of each other along the axis of commodity 1, that is if \( x \sim y \) then \( x + \alpha e_1 \sim y + \alpha e_1 \) for \( e_1 = (1, 0, 0, 0, \ldots, 0) \) and any scalar \( \alpha > 0 \), and
  \item good 1 is desirable, that is \( x + \alpha e_1 \succ x \) for all \( x \) and \( \alpha > 0 \).
\end{enumerate}

Next continuity of preferences:

\textbf{Definition 11.} The preference relation \( \succeq \) on \( X \) is \textbf{continuous} if its upper and lower contour sets are closed.

Rational continuous preferences can be represented by a real-valued function which we call a \textbf{utility function}. This means there is a function \( U : X \to R \) such that \( U(y) \geq U(x) \) iff \( y \succeq x \).
Proposition 5. If the preference ordering $\succeq$ on $X$ is rational and continuous then there exists a continuous utility function that represents this ordering.

This means that the utility function numbers indifference sets in such a way that the numbering is higher for more preferred sets. Note that there are in general very many ways of doing this so that the utility representation is not unique. Preferences are in some sense more fundamental than the utility representation. And this is not always possible: continuity of the upper contour sets matters, as the next example shows.

An interesting case to think about and use to test your understanding is that of lexicographic preferences. To see what these are take the case of $\mathbb{R}^2$ so all points we look at are 2-vectors. Then define $\succeq$ by $x \succeq y$ if either $x_1 > y_1$ or $x_1 = y_1$ & $x_2 \geq y_2$. This is like the alphabetical ordering. What do indifference curves look like in this case? Is there a continuous utility function?

The properties of preference orderings translate into properties of utility functions:

1. Monotonicity implies that $U(x) > U(y)$ if $x \gg y$.

2. Convexity of preferences implies that $U(x)$ is quasi-concave: $\{y : U(y) \geq U(x)\}$ is convex for all $x$.

Proposition 6. A continuous preference on $X$ is homothetic if and only if it admits a utility function that is homogeneous of degree one, i.e. $U(\alpha x) = \alpha U(x) \forall \alpha > 0$.

Proof. If $\succeq$ is homothetic then $I(x) = \{y : y \sim x\}$ and $I(\alpha x) = \{y : y \sim \alpha x\} = \alpha I(x)$. So there is a utility function for which $U(\alpha x) = \alpha U(x)$ (in effect we number the indifference surfaces by $\alpha$). Now conversely suppose there is such a function. Then it’s obvious that the preference is homothetic.

Proposition 7. Note that if $U(x)$ represents $\succeq$ then so does $\phi(U(x))$ for any increasing function $\phi : \mathbb{R} \to \mathbb{R}$. So $U$ is unique only up to an order-preserving transformation, i.e. it is ordinal.

Proof. We show that the indifference curves of the two functions are the same. Look at $I_U(x) = \{y : U(y) = U(x)\}$ & $I_\phi(x) = \{y : \phi(U(y)) = \phi(U(x))\}$. If $U(x) = U(y) \Rightarrow \phi(U(x)) = \phi(U(y))$ so $I_U \subseteq I_\phi$. And if $\phi(U(x)) = \phi(U(y))$ then $U(x), U(y)$ are in the same level set for $\phi : \mathbb{R}^1 \to \mathbb{R}^1 \Rightarrow U(x) = U(y) \Rightarrow I_\phi \subseteq I_U$. 

10
This means that the scale of the utility function is arbitrary, and that there is no meaning to utility differences - we can’t say that “1 is preferred to 2 by more than 3 is preferred to 4” because preference differences are not meaningful: we can only rank options, we can’t rank their differences. We can’t measure preference intensity.

**Proposition 8.** A quasi-linear preference can be represented by a utility function of the form $U(x) = x_1 + \theta(x_2, \ldots, x_N)$.

In the special case of $N = 2$ then $U(x) = x_1 + \theta(x_2)$ and indifference curves are all parallel but displaced along the horizontal axis. The marginal rate of substitution between the two goods is independent of the consumption of the first good and depends only on that of the second.

## 2 Utility Maximization Problem (UMP)

Consider the utility maximization problem $\max U(x), p.x = W$. We assume from now on that there is a unique solution to this problem. A sufficient condition for this is that upper contour sets are strictly convex. This of course implies that we have demand functions rather than correspondences.

**Proposition 9.** Suppose that $U(x)$ is a continuous utility function representing a non-satiated preference ordering $\succeq$ on the consumption set $X = \mathbb{R}_+^N$. The demand function $x(p, W)$ has the following properties:

1. Homogeneous of degree zero in $(p, W)$
2. Walras’ Law: $p.x = W$ for all $x \in x(p, W)$

**Proof.** Clearly $x(p, W) = x(\alpha p, \alpha W)$ proving 1. Point 2 follows from non-satiation.

### 2.1 Digression on constrained maximization

We can turn the constrained maximization problem $\max U(x), p.x = W$ into an unconstrained problem by introducing the Lagrangian

$$L = U(x) + \lambda[p.x - W]$$  \hspace{1cm} (2.1)

where $\lambda \in \mathbb{R}_+^1$ is a Lagrange multiplier. We assume the function $U$ to be concave: in this case
Proposition 10. If there exists $\lambda^* \geq 0$ such that $(x^*, \lambda^*)$ form a saddle point of the Lagrangian then $x^*$ solves the constrained utility maximization problem.

Proof. Suppose $(x^*, \lambda^*)$ satisfy:

$$ x^* \max L(x, \lambda^*) $$
$$ \lambda^* \min L(x^*, \lambda) $$

which means that they form a saddle-point of $L$. Then $L(x^*, \lambda^*) = U(x^*) + \lambda^*(p.x^* - W) = U(x^*)$ because $\partial L/\partial \lambda = (p.x^* - W) = 0$. Hence at $x^*$ the budget constraint is satisfied. Now note that $L(x^*, \lambda^*) \geq L(x, \lambda^*) \forall x$ which implies that $U(x^*) + \lambda^*(p.x^* - W) \geq U(x) + \lambda^*(p.x - W) \forall x$ and as $p.x^* = W$ we have $U(x^*) \geq U(x) + \lambda^*(p.x - W) \forall x$ and so in particular for any $x$ satisfying $p.x = W$. Hence $x^*$ is a constrained maximum.

2.2 UMP solutions

Necessary conditions for $x^*$ to be a solution to the UMP are that there exist a Lagrange multiplier $\lambda$ such that:

$$ \frac{\partial U(x^*)}{\partial x_n} \leq \lambda p_n, \text{ with equality if } x_n^* > 0 $$

So if we are at an interior optimum then

$$ \frac{\partial U(x^*)}{\partial x_n} = \lambda p_n $$

In matrix notation, letting $\nabla U(x) = (\partial U/\partial x_1, \ldots, \partial U/\partial x_N)$, we have

$$ \nabla U(x^*) \leq \lambda p, \ x^*. [\nabla U(x) - \lambda p] = 0 $$

This tells us that at an interior solution the gradient vector of the utility function must be proportional to the price vector, or more intuitively marginal rates of substitution between commodities must be equal to their price ratios:

$$ \frac{\partial U(x^*)/\partial x_k}{\partial U(x^*)/\partial x_j} = p_k/p_j $$

Note that the value of the Lagrange multiplier $\lambda$ gives us the marginal utility of wealth, i.e. the rate at which utility increases with wealth if the budget constraint is relaxed slightly. To see this note that if $W$ changes by $\Delta W$ then the resulting change in utility is

$$ \Delta U = \sum \frac{\partial U}{\partial x_n} \Delta x_n = \lambda \sum p_n \Delta x_n = \lambda \Delta W $$
**Definition 12.** Let $V(p, W) = U(x^*)$ for any $x^* \in x(p, W)$. Then $V(p, W) : R^{N+1} \rightarrow R$ is called the indirect utility function. $V(p, W) = U(x(p, W))$

**Proposition 11.** If $U$ is continuous and represents a locally non-satiated preference ordering then the indirect utility function $V(p, W)$ is

1. Homogeneous of degree zero
2. Strictly increasing in $W$ and non-increasing in $p_n$ for any $n$
3. Quasi-convex, i.e. $\{p, W : V(p, W) \leq \hat{V}\}$ is convex for any $\hat{V}$
5. $\frac{\partial V}{\partial W} = \lambda$
6. $\frac{\partial V}{\partial p_n} = -\lambda x_n (p, W)$

**Proof.** Point 6: $\frac{\partial V}{\partial p_n} = \sum_j \frac{\partial U}{\partial x_j} \frac{\partial x_j}{\partial p_n} = \sum_j \lambda p_j \frac{\partial x_j}{\partial p_n}$. But $\sum_j x_j p_j = W \Rightarrow x_n + \sum_j p_j \frac{\partial x_j}{\partial p_n} = 0$. So $\sum_j \lambda p_j \frac{\partial x_j}{\partial p_n} = -\lambda x_n$. \hfill \Box

**Example 1: Cobb-Douglas Utility**

$$Max_{x_1, x_2} x_1^\alpha x_2^{1-\alpha}, \quad p_1 x_1 + p_2 x_2 \leq W$$

As $U$ is monotone in on both arguments, we know that a solution will be on the budget hyperplane. Write out a Lagrangian

$$L = x_1^\alpha x_2^{1-\alpha} + \lambda [W - p_1 x_1 - p_2 x_2]$$

$$\alpha x_1^{\alpha-1} x_2^{1-\alpha} = \lambda p_1$$

$$(1 - \alpha) x_1^\alpha x_2^{-\alpha} = \lambda p_2$$

Dividing the two FOCs we get

$$\frac{\alpha}{(1 - \alpha)} \frac{x_2}{x_1} = \frac{p_1}{p_2} \Rightarrow x_2 = x_1 \frac{p_1}{p_2} \frac{(1 - \alpha)}{\alpha}$$

From the budget constraint

$$p_1 x_1 + p_2 x_1 \frac{p_1}{p_2} \frac{(1 - \alpha)}{\alpha} = W$$

Hence

$$x_1 = W \frac{\alpha}{p_1}, \quad x_2 = W \frac{1 - \alpha}{p_2}$$
From these we can work out the income and price elasticities:

\[ PED_{1,1} = PED_{2,2} = -1, PED_{1,2} = PED_{2,1} = 0, IED_1 = IED_2 = 1 \]

Note that the cross PEDs are zero, which is obvious when you rewrite the utility function as \( \alpha \ln x_1 + (1 - \alpha) \ln x_2 \). We can work with the log of the utility function as it is an increasing transformation and so preserves the underlying ordering, i.e. the upper contour sets are the same as before. It is often interesting to look at expenditure shares as a function of price and income. In this case we have

\[ S_1 = \frac{p_1 x_1}{W} = \alpha, \quad S_2 = \frac{p_2 x_2}{W} = 1 - \alpha \]

so the expenditure shares are just the exponents of the consumption levels in the Cobb-Douglas case.

Indirect utility function: by definition \( V(p, W) = U(x^*) \). From above

\[ x_1 = W \frac{\alpha}{p_1}, x_2 = W \frac{1 - \alpha}{p_2} \] so \( V(p, W) = W \left( \frac{\alpha}{p_1} \right)^\alpha \left( \frac{1 - \alpha}{p_2} \right)^{1-\alpha} \) which is linear in \( W \).

**Example 2: Linear Utility**

\[
\max_x \{ ax_1 + bx_2 \} \text{ s.t. } p_1 x_1 + p_2 x_2 = W
\]

There exists \( \lambda \) such that

\[
\frac{\partial U}{\partial x_i} \leq \lambda p_i, \quad = \text{ if } x_i > 0, i = 1, 2
\]

So

\[
a \leq \lambda p_1, \quad = \text{ if } x_1 > 0; \quad b \leq \lambda p_2, \quad = \text{ if } x_2 > 0
\]

So if \( x_1, x_2 > 0 \) then \( a/b = p_1/p_2 \). Otherwise either \( x_1 > 0 \) and \( x_2 = 0 \) or vice versa.

**Example 3: Fixed Coefficients**

Utility is \( U(x) = \min \{ ax_1, bx_2 \} \) so the problem is

\[
\max_x \min \{ ax_1, bx_2 \}, \quad p_1 x_1 + p_2 x_2 = W
\]

We know that to avoid wasting either good we need \( ax_1 = bx_2 \) so we want to \( \max x_1 \text{ with } x_2 = x_1 a/b \). So \( p_1 x_1 + p_2 x_1 \frac{a}{b} = W \) and

\[
x_1 = \frac{bW}{p_1 a + p_2 b}, \quad x_2 = \frac{aW}{p_1 a + p_2 b}
\]
In this case the indirect utility function is

\[ V(p_1, p_2, W) = ax_1 = bx_2 = \frac{abW}{p_1a + p_2b} \]

**Example 4: Constant Elasticity of Substitution**

The problem is \( \text{Max} \{x_1^\rho + x_2^\rho\}^{1/\rho} \), \( p_1x_1 + p_2x_2 = W \). Here \( 1/(1 - \rho) \) is the elasticity of substitution between \( x_1 \) and \( x_2 \). For \( \rho = 1 \) the function is linear, and for \( \rho \to -\infty \) it shows fixed coefficients. When \( \rho = 0 \) it is Cobb-Douglas. Lagrangian is

\[ L = \{x_1^\rho + x_2^\rho\}^{1/\rho} + \lambda (W - p_1x_1 - p_2x_2) \]

FOCs are

\[ \frac{1}{\rho} \{x_1^\rho + x_2^\rho\}^{\frac{1}{\rho} - 1} \rho x_i^{\rho - 1} = \lambda p_i, \quad i = 1, 2, \quad \text{so} \quad \frac{x_1}{x_2} = \left( \frac{p_1}{p_2} \right)^{\frac{1}{\rho - 1}} \]

Using the budget constraint

\[ x_2 \left( \frac{p_1}{p_2} \right)^{\frac{1}{\rho - 1}} p_1 + x_2p_2 = x_2 \left\{ p_1^{\frac{1}{\rho - 1} + \frac{1}{\rho - 1}} + p_2^{\frac{1}{\rho - 1}} \right\} = W \]

so

\[ x_2p_2 \left\{ p_1^{\frac{1}{\rho - 1}} + p_2^{\frac{1}{\rho - 1}} + 1 \right\} = W \]

Let \( r = \rho / (\rho - 1) \). Then

\[ x_2p_2 \left\{ p_1^{r} + p_2^{r} + 1 \right\} = x_2p_2 \left\{ \frac{p_1^{r} + p_2^{r}}{p_2^{r}} \right\} = W \]

so finally

\[ x_2 = W \frac{p_2^{r - 1}}{p_1^{r} + p_2^{r}}, \quad x_1 = W \frac{p_1^{r - 1}}{p_1^{r} + p_2^{r}} \]

Indirect utility is

\[ V(p_1, p_2, W) = \left\{ \left( \frac{Wp_1^{r - 1}}{p_1^{r} + p_2^{r}} \right)^{\rho} + \left( \frac{Wp_2^{r - 1}}{p_1^{r} + p_2^{r}} \right)^{\rho} \right\}^{1/\rho} = W \left\{ \frac{p_1^{r} + p_2^{r}}{(p_1^{r} + p_2^{r})^{\rho}} \right\}^{1/\rho} = W (p_1^{r} + p_2^{r})^{-\frac{1}{\rho}} \]
3 Expenditure Minimization Problem (EMP)

Consider the problem

\[
\min_{x \geq 0} p.x, \text{ s.t. } U(x) \geq \hat{U}
\]

This is the EMP, the dual to (opposite of) the UMP \(\max \ x, \text{ s.t. } U(x) \leq W.\)

Here we seek to minimize expenditure subject to not falling below a specified utility level: in the UMP we maximize utility subject to not exceeding a specified expenditure level.

**Proposition 12.** Suppose \(U\) is continuous and represents a locally non-satiated preference on \(\mathbb{R}^N_+\) and that the price vector \(p \gg 0\). Then

1. If \(x^*\) is optimal in the UMP when wealth is \(W > 0\) then \(x^*\) is also
   optimal in the EMP when the required utility level is \(\hat{U} = U(x^*)\). Moreover
   the minimized expenditure level in the EMP is exactly \(W\).

2. If \(x^*\) is optimal in the EMP when the required utility level \(\hat{U} > U(0)\),
   then \(x^*\) is optimal in the UMP with wealth \(p.x^*\). Moreover the maximized
   utility level in this UMP is exactly \(\hat{U}\).

**Definition 13.** The expenditure function \(e(p,U)\) is the solution to the
EMP problem for prices \(p\) and required utility \(U\).

Its properties are similar to those of the indirect utility function:

**Proposition 13.** Suppose \(U\) is continuous and represents a locally non-satiated preference on \(\mathbb{R}^N_+\). Then the expenditure function \(e(p,U) : \mathbb{R}^N \to \mathbb{R}\) is

1. Homogeneous of degree one in \(p\)
2. Strictly increasing in \(U\) and non-decreasing in \(p_n, \forall n\).
3. Concave in \(p\).
4. Continuous in \(p\) and \(U\).

**Example: Cobb-Douglas**

To compute the expenditure function for the Cobb-Douglas utility considered in earlier examples we will take a monotone transform of this function by taking logs: \(U(x) = \alpha \ln x_1 + (1 - \alpha) \ln x_2\). The problem is

\[
\min_x (p_1x_1 + p_2x_2) \text{ s.t. } \alpha \ln x_1 + (1 - \alpha) \ln x_2 \geq U
\]

\(^4\)We get the dual of a constrained maximization problem by interchanging the objective and constraint and minimizing rather than maximizing.
and we assume the constraint holds with equality by non-satiation.

\[
L = p_1 x_1 + p_2 x_2 + \lambda (U - \alpha \ln x_1 - (1 - \alpha) \ln x_2)
\]

\[
p_1 = \frac{\alpha}{x_1}, \quad p_2 = \frac{1 - \alpha}{x_2}, \quad x_1 = x_2 \frac{p_2}{p_1} \frac{\alpha}{1 - \alpha}
\]

Using the constraint

\[
\alpha \ln x_2 + \alpha \ln \left( \frac{p_2 \alpha}{p_1 1 - \alpha} \right) + (1 - \alpha) \ln x_2 = U
\]

\[
\ln x_2 = U - \alpha \ln \left( \frac{p_2 \alpha}{p_1 1 - \alpha} \right)
\]

\[
x_1 = e^U \left( \frac{p_2}{p_1} \right)^{1 - \alpha} \left( \frac{\alpha}{1 - \alpha} \right)^{\alpha}, \quad x_2 = e^U \left( \frac{p_1}{p_2} \right)^{\alpha} \left( \frac{1 - \alpha}{\alpha} \right)^{\alpha}
\]

The expenditure function is thus

\[
e^U \left\{ p_1 \left( \frac{p_2}{p_1} \right)^{1 - \alpha} \left( \frac{\alpha}{1 - \alpha} \right)^{1 - \alpha} + p_2 \left( \frac{p_1}{p_2} \right)^{\alpha} \left( \frac{1 - \alpha}{\alpha} \right)^{\alpha} \right\} = e^U \left( \frac{\alpha}{1 - \alpha} \right)^{-\alpha} \left( \frac{p_2}{p_1} \right)^{-\alpha} \frac{p_2}{1 - \alpha}
\]

\[
= e^U \left( \frac{p_1}{\alpha} \right)^{\alpha} \left( \frac{p_2}{1 - \alpha} \right)^{1 - \alpha}
\]

**Example: Fixed Coefficients**

We know that

\[V((p_1, p_2, W) = \frac{abW}{p_1b + p_2a} \Rightarrow W = Expenditure = U \frac{p_1a + p_2b}{ab}\]

**Example: CES**

We know that

\[U = V(p_1, p_2, W) = W (p_1^r + p_2^r)^{-\frac{1}{r}} \Rightarrow W = Expenditure = U (p_1^r + p_2^r)\]

**Definition 14.** The optimal commodity vector in the EMP, denoted \( h(p, U) : \mathbb{R}^N \rightarrow \mathbb{R}^N \), is known as the **Hicksian or compensated demand function**.

The adjective “compensated” indicates that this is demand as a function of prices with the utility level constant, which means that as prices change wealth must be altered too to allow the consumer to maintain the same utility level.
Proposition 14. Suppose $U$ is continuous and represents a locally non-satiated preference $\succeq$ on $\mathbb{R}_+^N$. Then for any $p \gg 0$ the Hicksian demand function $h(p,U)$ is

1. homogeneous of degree zero in $p$
2. $U(h(p,U)) = U$ for any $p$

Next we establish a connection between the expenditure function and the compensated (Hicksian) demand function. Note that $e(p,U) = h(p,U) \cdot p$

Proposition 15. Let $U$ be a continuous utility function representing a locally non-satiated strictly convex preference relation $\succeq$ on $X$. For all $p$ and $U$ the Hicksian demand $h(p,U)$ is the derivative vector of the expenditure function $e(p,U)$ with respect to prices:

$$h(p,U) = \nabla_p e(p,U) \quad \text{or} \quad h_n(p,U) = \frac{\partial e(p,U)}{\partial p_n} \forall n$$

Proof. Assume for simplicity that $h(p,U) \gg 0$ and is differentiable. Using the chain rule we can write

$$\frac{\partial e(p,U)}{\partial p_n} = \frac{\partial}{\partial p_n} \sum_i h_i(p,U) p_i = \sum_i p_i \frac{\partial h_i}{\partial p_n} + h_n(p,U)$$

Using the FOCs this is

$$\frac{\partial e(p,U)}{\partial p_n} = \lambda \sum_i \frac{\partial h_i}{\partial p_n} \frac{\partial U}{\partial h_i} + h_n(p,U)$$

Note that $U = U(h(p,U)) \forall p$ so that $\frac{\partial h_i}{\partial p_n} \frac{\partial U}{\partial h_i} = 0$. This proves the proposition. \qed

So

- the compensated demand function is the regular demand function evaluated at the same prices and at the wealth level required to reach the specified utility level, and

- the regular demand function is the compensated demand function evaluated at the same prices and at the indirect utility function evaluated at the same prices and wealth.
The next proposition states that the effect of a price change on demand can be decomposed into two parts, one due to the price change alone at constant welfare level and the other due to the change in real income resulting from a price change. These are called the substitution and income effects respectively.

**Proposition 16.** (The Slutsky equation) Assume $U$ is a continuous utility function representing a locally non-satiated strictly convex preference relation $\succeq$ defined on $X$. Then for all $(p, W)$ and $U = V(p, W)$ we have

$$\frac{\partial x_n(p, W)}{\partial p_k} = \frac{\partial h_n(p, W)}{\partial p_k} - \frac{\partial x_n(p, W)}{\partial W} x_k(p, W) \forall n, k$$

or equivalently in matrix notation

$$D_p x(p, W) = D_p h(p, U) - D_W x(p, W) x(p, W)^T$$

**Proof.** Consider a consumer facing the price-wealth pair $\bar{p}, \bar{W}$ and attaining utility level $\bar{U}$. Her wealth level must satisfy $\bar{W} = e(\bar{p}, \bar{U})$. For all $(p, U)$ we have $h_n(p, U) = x_n(p, e(p, U))$. Differentiating this with respect to $p_k$ and evaluating at $(\bar{p}, \bar{U})$ gives

$$\frac{\partial h_n(\bar{p}, \bar{U})}{\partial p_k} = \frac{\partial x_n(\bar{p}, e(\bar{p}, \bar{U}))}{\partial p_k} + \frac{\partial x_n(\bar{p}, e(\bar{p}, \bar{U}))}{\partial W} \frac{\partial e(\bar{p}, \bar{U})}{\partial p_k} \tag{3.1}$$

This yields from proposition 15

$$\frac{\partial h_n(\bar{p}, \bar{U})}{\partial p_k} = \frac{\partial x_n(\bar{p}, e(\bar{p}, \bar{U}))}{\partial p_k} + \frac{\partial x_n(\bar{p}, e(\bar{p}, \bar{U}))}{\partial W} h_k(\bar{p}, \bar{U})$$

Finally since $\bar{W} = e(\bar{p}, \bar{U})$ and $h_k(\bar{p}, \bar{U}) = x_k(\bar{p}, e(\bar{p}, \bar{U})) = x_k(\bar{p}, \bar{W})$ we have

$$\frac{\partial h_n(\bar{p}, \bar{U})}{\partial p_k} = \frac{\partial x_n(\bar{p}, \bar{W})}{\partial p_k} + \frac{\partial x_n(\bar{p}, \bar{W})}{\partial W} x_k(\bar{p}, \bar{W})$$

Simple rearrangement gives

$$\frac{\partial x_n(\bar{p}, \bar{W})}{\partial p_k} = \frac{\partial h_n(\bar{p}, \bar{U})}{\partial p_k} - \frac{\partial x_n(\bar{p}, \bar{W})}{\partial W} x_k(\bar{p}, \bar{W})$$

19
Now let \( k = n \) so that we are looking at own price effects:

\[
\frac{\partial x_n(\bar{p}, \bar{W})}{\partial p_n} = \frac{\partial h_n(\bar{p}, \bar{U})}{\partial p_n} - \frac{\partial x_n(\bar{p}, \bar{W})}{\partial W} x_n(\bar{p}, \bar{W})
\]

This allows an intuitive interpretation. It gives a decomposition of the effect of a price change on Marshallian (conventional) demand into two components. A price change has two effects: it changes the relative prices of commodities, and it makes the consumer better or worse off, that is raises or lowers her real income. The first of these is the substitution effect, the effect of a price change alone, keeping utility constant: \( \frac{\partial h_n(\bar{p}, \bar{U})}{\partial p_n} \). The second is an income effect, the effect of the price change on real income, \( \frac{\partial x_n(\bar{p}, \bar{W})}{\partial W} x_n(\bar{p}, \bar{W}) \). We can also express this in terms of elasticities by multiplying through by \( \frac{p_n}{x_n} \):

\[
\frac{\partial x_n(\bar{p}, \bar{W})}{\partial p_n} \frac{p_n}{x_n} = \frac{\partial h_n(\bar{p}, \bar{U})}{\partial p_n} \frac{p_n}{x_n} - \frac{\partial x_n(\bar{p}, \bar{W})}{\partial W} \frac{W}{x_n} \frac{p_n x_n}{W}
\]

which states that the elasticity of regular (Marshallian) demand equals that of compensated (Hicksian) demand minus the income elasticity of demand times the expenditure share.

### 3.1 Using First Order Conditions

Consider a consumer problem

\[ \text{Max} \, x \, U(x \mid y), \, p.x = W \]

where \( y \) is a parameter that affects preferences - for example, temperature affects preferences for drinks and clothes. For the two-good case the FOC are

\[
\frac{U_1(x \mid y)}{U_2(x \mid y)} - \frac{p_1}{p_2} = F(x, y, p) = 0
\]

We can use the implicit function theorem to get the impact of a change in the parameter \( y \) on the demand for say \( x_1 \).

\[
\frac{\partial x_1}{\partial y} = -\frac{F_y}{F_{x_1}} = \frac{U_2 U_{1y} - U_1 U_{2y}}{U_2 U_{11} - U_1 U_{21}}
\]

and for a particular functional forms we can determine the sign of this.
3.2 Derivation of Slutsky Equation for labor supply

Another example is the supply curve for labor as a function of the wage rate. Let utility be $U(y, L)$ where $y$ is consumption and $L$ is leisure, so $U$ is increasing in both. Income is given by $y = (K - L)w + A$ where $w$ is the wage rate and $K$ is the number of hours available for work and leisure. $A$ is the agent’s non-labor income. The consumer problem is

$$\text{Max}_L U(y, L), \ y + wL = wK + A$$

and we let $S = wK + A$ be the total income the agent could earn if he devoted all his time to work. Substituting for $y$ gives $U(w(K - L + A, L))$ and maximizing with respect to $L$ gives as FOCs This is his wealth: unlike in previous cases it depends on the price $w$. The FOCs are

$$-Uw + U_L = 0 \text{ or } U_L = wU_y$$

From this we get the Marshallian demand function $x_l(p, w, S)$ for leisure and by solving the expenditure minimization problem we get the Hicksian or compensated demands $h_l(p, w, U)$ for leisure. Using the Slutsky equation we can write

$$\frac{\partial x_l}{\partial w} = \frac{\partial h_l}{\partial w} - \left( \frac{\partial h_l}{\partial S} \right) L$$

Note that $x_l(p, w, S) = x_l(p, w, wK + A)$ from which

$$\frac{dx_l^*}{dw} = \frac{\partial x_l}{\partial w} + K \frac{\partial x_l}{\partial S}$$

Note that $x_l$ depends on $w$ via two of its arguments so we need to take this into account when differentiating with respect to $w$. I have used $dx_l^*/dw$ to stand for the derivative taking this into account. The standard derivation of the Slutsky equation does not take this into account.

Substituting this into the Slutsky equation we have

$$\frac{\partial x_l}{\partial w} = \frac{dx_l^*}{dw} - K \frac{\partial x_l}{\partial S} = \frac{\partial h_l}{\partial w} - L \frac{\partial x_l}{\partial S}$$

so

$$\frac{dx_l^*}{dw} = \frac{\partial h_l}{\partial w} + \left( \frac{\partial h_l}{\partial S} \right) (K - L)$$

Here the first term is negative as it is the own price effect: an increase in the wage rate makes leisure more expensive and reduces consumption of leisure (substitution effect), while if leisure is a normal good the second term is positive representing the income effect of a wage change.
3.2.1 Example

We illustrate this with a Cobb-Douglas utility function. \( U(y, L) = y^\alpha L^{1-\alpha} = w^\alpha (24 - L)^{\alpha} L^{1-\alpha} \) so the FOC is

\[
\frac{\partial U}{\partial L} = -w^\alpha (24 - L)^{\alpha-1} L^{1-\alpha} + w^\alpha (1 - \alpha) (24 - L)^{\alpha} L^{-\alpha} = 0
\]

which implies

\[
\frac{L}{24 - L} = \frac{1 - \alpha}{\alpha}
\]

so the ratio of leisure to work equals the ratio of the exponents of leisure and work. Labor supply is independent of the wage rate: income effects exactly offset substitution effects. In the first problem set you will investigate whether this is true for a CES utility function.

Now we switch to an application in the cost-benefit area:

3.3 Welfare Evaluation of Economic Changes

An issue we can investigate is the effect of a price change on consumer welfare - this could be the result of a policy measure such as a tax. Suppose prices change from \( p^0 \) to \( p^1 \): by how much is the consumer better or worse off? In principle the indirect utility function can tell us this: it goes from \( V(p^0, W) \) to \( V(p^1, W) \) and we just need to evaluate the difference \( V(p^1, W) - V(p^0, W) \). In general this number is hard to interpret as we don’t know what the units are: they are utility and the utility function is unique only up to an order-preserving transformation. However starting from an indirect utility function \( V(p, W) \) and an arbitrary strictly positive price vector \( \bar{p} \) we can consider the expenditure function \( e(\bar{p}, V(p, W)) \). This function gives the wealth needed to reach the utility level \( V(p, W) \) when prices are \( \bar{p} \). This is strictly increasing in the value of \( V \), and so is an indirect utility function, but giving answers in dollars rather than utility. So

\[
e(\bar{p}, V(p^1, W)) - e(\bar{p}, V(p^0, W))
\]

provides a measure of the welfare change when the price changes from \( p^0 \rightarrow p^1 \): the measure is in dollars. We can do this for any price vector \( \bar{p} \), and natural choices are \( p^0 \) and \( p^1 \). Associated with these choices are the equivalent variation \( EV \) and compensating variation \( CV \), introduced by Hicks.
Let $U^{0} = V (p^{0}, W)$, $U^{1} = V (p^{1}, W)$ and note that as we are holding income constant $e (p^{0}, U^{0}) = e (p^{1}, U^{1}) = W$. Define

$$EV (p^{0}, p^{1}, W) = e (p^{0}, U^{1}) - e (p^{0}, U^{0}) = e (p^{0}, U^{1}) - W$$

$$CV (p^{0}, p^{1}, W) = e (p^{1}, U^{1}) - e (p^{1}, U^{0}) = W - e (p^{1}, U^{0})$$

The interpretation of the EV is as follows. Imagine you ask the consumer, before the price change has occurred: “Would you rather have the price change from $p^{0}$ to $p^{1}$ or a cash payment of $x$?” The the EV is the cash payment at which the consumer is indifferent between the two. It is, for her, equivalent to the price change. (If the prices increase then the EV is negative.) Note that $e (p^{0}, U^{1})$ is the expenditure needed to achieve utility level $U^{1} = V (p^{1}, W)$, the level generated by the price change, at the prices $p^{0}$. So EV is therefore the extra wealth needed to compensate for the price change. We can also write the EV as follows:

$$V (p^{0}, W + EV) = U^{1}$$

The CV answers a different question: suppose the price change has already occurred, and we ask how much it would take to compensate the consumer for it and bring her back to her original welfare level at the new prices. The question is: “How much do I have to pay you to compensate you for the price change?” Note that $e (p^{1}, U^{0})$ tells us how much it costs to attain the initial welfare level at the new prices $p^{1}$. So the CV is the change in wealth needed to compensate for the new prices and get back to the initial state. We can write the CV as

$$V (p^{1}, W - CV) = U^{0}$$

In each case we are trying to restore the consumer to the original welfare level but in the EV case using the initial prices and in the CV case using the final prices. These two clearly in general give different answers, but they will nevertheless give the same ranking of alternatives - the consumer is better off under $p^{1}$ if and only if these measures are both positive. For normal goods (IED > 0) we have that $EV \geq CV$ and in the case of quasi-linear preferences we have $EV = CV$. To see this consider the two-good case with just the price of good 1 changing so that $p^{0}_{1} \neq p^{1}_{1}$, $p^{0}_{2} = p^{1}_{2}$. We can express the EV in terms of the Hicksian or compensated demand.
Recall that \( w = e(p^0, U^0) = e(p^1, U^1) \) and that (by proposition 15) \( h_1(p, U) = \partial e(p, U) / \partial p_1 \). Hence we can write

\[
EV(p^0, p^1, W) = e(p^0, U^1) - W = e(p^0, U^1) - e(p^1, U^1) = \int_{p_1^0}^{p_1^1} h_1(p_1, p_2, U^1) \, dp_1
\]

so the EV can be represented by the area between \( p_1^1 \) and \( p_1^0 \) to the left of the Hicksian demand curve for good 1 associated with utility level \( U^1 \). Similarly the CV can be expressed as

\[
CV(p^0, p^1, W) = \int_{p_1^0}^{p_1^1} h_1(p_1, p_2, U^0)
\]

which is the area between the two prices to the left of the Hicksian demand curve corresponding to utility level \( U^0 \). Clearly

\[
EV(p^0, p^1, W) - CV(p^0, p^1, W) = \int_{p_1^0}^{p_1^1} \{ h_1(p_1, p_2, U^1) - h_1(p_1, p_2, U^0) \}
\]

which is zero, so that \( EV = CV \), if the Hicksian demand for good 1 is independent of the utility level \( U \). This is the case for quasi-linear preferences. To see this assume \( U(x) = f(x_1) + x_2 \) and recall that the Hicksian demand is the solution to the expenditure minimization problem

\[
Min \{ p_1x_1 + p_2x_2 \}, \ f(x_1) + x_2 = \hat{U}
\]

This reduces to \( Min_{x_1, p_1x_1 + p_2} \left[ \hat{U} - f(x_1) \right] \) and the solution to this is \( x_1^* = (f')^{-1} \left( \frac{p_1}{p_2} \right), \ X_2^* = \hat{U} - f(x_1^*) \). Hence the expenditure function is

\[
e(p_1, p_2, \hat{U}) = p_1 \left( f' \right)^{-1} \left( \frac{p_1}{p_2} \right) + p_2 \left( \hat{U} - f \left( (f')^{-1} \left( \frac{p_1}{p_2} \right) \right) \right)
\]

and the derivative of this with respect to \( p_1 \) is independent of the utility level. In this case the EV and CV are equal and are both equal to the conventional Marshallian consumer surplus, the area between the two prices to the left of the regular demand curve.
3.4 Deadweight loss from commodity taxation.

A standard question in public finance is: which way of raising government revenue reduces consumer welfare least? We will use the EV-CV machinery to compare the costs of raising money by commodity taxes with the cost of raising it by a lump-sum tax, a subtraction from wealth.

Consider a two commodity world with initial prices $p_0^0, p_0^1$, where a tax of $t$ per unit is levied on the first good, so that its price changes to $p_1^1 = p_0^1 + t$. Revenue raised is $T = tx_1 (p_1^1, W)$. The alternative is a tax of $T$ on wealth, reducing this to $W - T$.

The consumer is worse off under the commodity tax if $EV (p_1^1, p_0^1, W)$, which is negative, is more negative than $-T$, that is $EV < -T$ or $0 < -T - EV$.

In terms of expenditure functions she is worse off under the commodity tax if $W - T > e (p_0^0, U_1)$ or $W - T - e (p_0^0, U_1) > 0$, that is if the wealth she has after the lump-sum tax exceeds the wealth needed at prices $p_0^0$ to generate the utility she gets under the commodity tax. We can equate these two criteria:

$$-T - EV (p_0^0, p_1^1, W) = W - T - e (p_0^0, U_1)$$

This is called the **deadweight loss from commodity taxation**. It measures how much worse off the consumer is because of the use of commodity rather than lump-sum taxation. We can write this in terms of the Hicksian or compensated demand curve at utility level $U_1$.

$$-T - EV (p_0^0, p_1^1, W) = e (p_1^1, U_1) - e (p_0^0, U_1) - T$$

Note that $e (p_1^1, U_1) - e (p_0^0, U_1) = EV$ and $T = tx_1 (p_1^1, W) = th_1 (p_1^1, U_1)$

$$= \int_{p_0^0}^{p_0^0 + t} h_1 (p_1, p_2, U_1) \, dp_1 - th_1 (p_1^0 + t, p_2, U_1)$$

$$= \int_{p_0^0}^{p_0^0 + t} \left\{ h_1 (p_1, p_2, U_1) \, dp_1 - h_1 (p_1^0 + t, p_2, U_1) \right\} \, dp_1$$

as $p_0^0$ is constant independent of $t$. Because $h_1$ is non-increasing in $p_1$, this expression is non-negative, and is strictly positive if $h_1$ is strictly decreasing in $p_1$. 

25
4 First Problem Set

1. Compute demand functions for goods 1 and 2 when the utility functions are (1) \( U(x_1,x_2) = x_1^3 x_2^5 \) and (2) \( U(x_1,x_2) = 3x_1^{1/2} + x_2^{1/2} \).

2. Show that if \( x(p,W) \) is homogeneous of degree one with respect to \( W \), i.e. \( x(p,aW) = ax(p,W) \forall a > 0 \) and satisfies Walras’ Law, then \( \varepsilon_{l,W} = 1 \) for every \( l \) where \( \varepsilon_{l,W} \) is the elasticity of demand for good \( l \) with respect to wealth. Can you say something about \( D_W(x,p) \) and the form of the Engel functions and curves in this case?

3. Show that the elasticity of demand for good \( l \) with respect to price \( p_k \), \( \varepsilon_{l,k} \), can be written as \( \varepsilon_{l,k} = d\ln(x_l(p,W))/d\ln p_k \). Derive a similar expression for \( \varepsilon_{l,W} \). Show that if we estimate the parameters \((a_0, a_1, a_2, b)\) in the equation \( \ln(x_l(p,W)) = a_0 + a_1 \ln p_1 + a_2 \ln p_2 + b \ln W \) these parameters provide estimates of the elasticities \( \varepsilon_{l,1}, \varepsilon_{l,2}, \varepsilon_{l,W} \).

4. Draw a convex preference relation that is locally non-satiated but not monotone.

5. Suppose that in a two commodity world the consumer’s utility function takes the form \( U(x) = [\alpha_1 x_1^{\rho} + \alpha_2 x_2^{\rho}]^{1/\rho} \), known as the CES or constant elasticity of substitution function. (A) Show that when \( \rho = 1 \), indifference curves are linear. (B) Show that as \( \rho \to 0 \), the utility function comes to represent the same preferences as the generalized Cobb-Douglas \( x_1^{a_1} x_2^{a_2} \). (C) Show that as \( \rho \to -\infty \), indifference curves become right angled, that is they become the indifference map of the Leontief function \( \min \{x_1, x_2\} \).

6. The elasticity of substitution between goods 1 and 2 is defined as

\[
\xi_{1,2} = \frac{-\frac{\partial [x_1/x_2]}{\partial [p_1/p_2]} p_1/p_2}{x_1/x_2}
\]

Show that for CES functions \( \xi_{1,2} = 1/(1 - \rho) \). What is this elasticity for the Linear, Cobb-Douglas and Leontief cases?

7. Consider a consumer who chooses his consumption bundle \( x_1, x_2, x_3, ..., x_n \) to maximize his utility \( U(x_1, ..., x_n) \) subject to the budget constraint \( \sum_k p_k x_k \leq Y \). Prove the “Engel Proposition” that the sum of the
products of each income elasticity with its budget share must equal one, i.e.

$$\sum_k \alpha_k E_k = 1$$

where \( E_k = \frac{\partial x_k}{\partial y} \frac{y}{x_k} \) and \( \alpha_k = \frac{p_k x_k}{Y} \).

8. For the utility function

$$U(x_1, x_2) = \beta_1 \ln (x_1 - \gamma_1) + \beta_2 \ln (x_2 - \gamma_2)$$

prove that demands \( x_i, i = 1, 2 \) satisfy the so-called “linear expenditure system”

$$p_i x_i = p_i \gamma_i + \beta_i \left( Y - \sum_i p_i \gamma_i \right) i = 1, 2$$

9. A consumer has the utility function \( \{y^\rho + L^\rho\}^{1/\rho} \) where \( y \) is income, \( L \) is leisure and the two are related by the budget \( y = w (24 - L) \) where \( w \) is the wage rate. What is the slope of the labor supply curve?
5 Preference Aggregation

The first issue to look at here is that of social indifference curves. Recall that for a rational preference relation \( \succeq \) on \( \mathbb{R}^N \) the preferred or indifferent set to the point \( x \) is \( PI(x) = \{ y : y \succeq x \} \). Let there be \( I \) individuals indexed by \( i \), each consuming \( x_i \). The for each person we have \( PI_i(x_i) = \{ y : y \succeq_i x_i \} \) where \( \succeq_i \) denotes the \( i \)-th individual’s preference ordering. Consider

\[
PI(x_1, \ldots, x_I) = \sum_i PI_i(x_i) = \left\{ Y : Y = \sum_i y_i, y_i \in PI_i(x_i) \forall i \right\}
\]

So this is the set of points that can be divided between agents so that each is in the preferred or indifferent set to \( x_i \). The boundary of this set is the social indifference curve SIC associated with \( (x_1, \ldots, x_I) \), \( SI(x_1, \ldots, x_I) \), and is the set of points that can be divided between agents so that each is on the indifference curve containing \( x_i \). An important question is: do these indifference curves define a rational preference - complete and transitive?

Complete is not an issue: transitive is.

**Proposition 17.** A sufficient condition for the SICs \( SI(x_1, \ldots, x_I) \) for all possible allocations \( x_1, \ldots, x_I \) to form a transitive preference is that all agents have identical and homothetic preferences. (Note: a necessary and sufficient condition is a bit weaker but not a lot: it is that each agent have a preference that is an affine translation of a given homothetic preference. See paper by me and Chichilnisky Jour Math Econ 1983 )

**Proof.** Recall that a preference is homothetic if and only if \( x \sim y \Rightarrow \alpha x \sim \alpha y \forall \alpha > 0 \). This implies that \( x \succeq y \Rightarrow \alpha x \succeq \alpha y \forall \alpha > 0 \). So if agent \( i \)'s preferences are homothetic then all preferred or indifferent sets are scalar multiples of each other: \( \exists \alpha > 0 : PI_i(x_i) = \alpha PI_i(y_i) \) for any \( x_i, y_i \). If all preferences are homothetic and identical there exists an individual \( k \) and an allocation \( x_k \) and scalars \( \alpha_j(x_j) \) such that

\[
PI_j(x_j) = \alpha_j(x_j) PI_k(x_k)
\]

Hence

\[
\sum_j PI_j(x_j) = PI_k(x_k) \sum_j \alpha_j(x_j) = \beta PI_k(x_k)
\]

so the social preferences are homothetic too and so transitive. \( \Box \)
So we can get a well-behaved aggregate preference or social preference from individual preferences only under restrictive conditions.

Next we look at the aggregation of demands rather than preferences. Each individual demand function depends on prices and income or wealth: \( d_i(p, W_i) \). Aggregate demand is \( D = \sum_i d_i(p, W_i) \). We can write this as \( D = D(p, W_1, \ldots, W_I) \) as the price is the same for all. An interesting question is: when does aggregate demand depend only on the total wealth/income rather than on the distribution? Formally, when do we have \( D = D(p, \sum_i W_i) \) so that the distribution does not affect aggregate demand? We are looking for conditions under which a change in the distribution of income/wealth will not change aggregate demand, that is any set of alterations in the wealth levels \( \Delta W_i, \sum_i \Delta W_i = 0 \), leads to no change in aggregate demand for any good. Letting \( d_{l,i} \) be \( i \)'s demand for good \( l \), this means

\[
\sum_i \frac{\partial d_{l,i}}{\partial W_i} \Delta W_i = 0 \forall l
\]

This is true if

\[
\frac{\partial d_{l,i}(p, W_i)}{\partial W_i} = \frac{\partial d_{l,k}(p, W_k)}{\partial W_k} = d'_l
\]

because in this case

\[
\sum_i \frac{\partial d_{l,i}}{\partial W_i} \Delta W_i = \sum_i d'_l \Delta W_i = 0
\]

This means that the effect of a change in wealth on demand is the same for all individuals and all wealth levels, which means that for a given price vector all wealth expansion paths are parallel. A sufficient condition for this is that all preferences are homothetic and identical. All preferences being quasi-linear with respect to the same good is also sufficient. As with preference aggregation, there is a weaker necessary and sufficient condition, but not a lot weaker. Here is an example for the quasi-linear case with two goods.

Let \( U_i(x_{i,1}, x_{i,2}) = x_{i,1} + f(x_{i,2}) \). If prices are \( p \) and wealth is \( W_i \) after using the budget constraint the utility maximization problem is

\[
Max_{x_{i,2}} \left\{ \frac{W_i}{p_1} - x_{i,2} \frac{p_2}{p_1} + f(x_{i,2}) \right\}
\]

The FOCs are

\[
x_{i,2} = (f')^{-1} \left( \frac{p_2}{p_1} \right), \quad x_{i,1} = \frac{W_i}{p_1} - \frac{p_2}{p_1} (f')^{-1} \left( \frac{p_2}{p_1} \right)
\]
Now consider a group of \( N \) consumers with such preferences and demands. Their aggregate demand is

\[
x_2 = N \left( f' \right)^{-1} \left( \frac{p_2}{p_1} \right), \quad x_1 = \sum W_i - \frac{p_2}{p_1} N \left( f' \right)^{-1} \left( \frac{p_2}{p_1} \right)
\]

which is of the same form as the individual demands and independent of the distribution of income.

Finally, return to the consumer’s expenditure minimization problem EMP:

\[
\text{Min}_{x_i} \{ p.x_i \}, \quad U_i(x_i) \geq \hat{U}_i
\]

Hicksian or compensated demand \( x^*_i(p, \hat{U}_i) = \text{argmin} \ p.x : U_i(x) \geq \hat{U}_i \).

We know that \( PX_i(x_i) = \{ y : U_i(y) \geq U_i(x_i) \} \) is a convex set, and so is \( \{ x : U_i(x) \geq \hat{U}_i \} \).

**Proposition 18.** \( x^*_i(p, \hat{U}_i) \) minimizes \( p.x \) over \( P_i(\hat{U}_i) = \{ x : U_i(x) \geq \hat{U}_i \} \)

for each \( i \) if and only if \( \sum_i x^*_i(p, \hat{U}_i) = X^* \left( p, \hat{U}_1, ..., \hat{U}_N \right) \) minimizes \( p.X \)

over \( \sum_i P_i(\hat{U}_i) \). In words, the order of set summation and minimization can be interchanged: the sum of the cost minima over the individual preferred or indifferent sets equals the minimum over the sum of the preferred or indifferent sets.

So although we can’t aggregate utility maxima we can aggregate expenditure minima, because there is no budget constraint and no income effect.

### 6 Preference aggregation and social choice

We have spoken of preference aggregation via the market (aggregation of demands) and via the summation of indifference curves, which turned out to be the same problem. Democratic systems also aggregate preferences - voting is a way of doing this and getting a social preference from a set of disparate individual preferences. Voting is an exercise in social choice. Let \( \Pi \) be the set of all possible individual preference orderings, \( \succeq_i \in \Pi \forall i \).

**Definition 15.** A social choice rule is a function \( \Phi : \Pi^N \rightarrow \Pi \) which associates with any \( N \)-tuple of individual preferences, each in \( \Pi \), a social preference, also in \( \Pi \).
All forms of voting are social choice rules - single non-transferable votes, transferable votes, proportional representation, etc. They all map a diverse set of individual preferences into a single social preference. Voting systems and other social choice systems run into the Condorcet Paradox. Consider a set of three people \{A, B, C\} with preferences over three alternatives \{α, β, γ\}. Let their rankings of these alternatives be as follows:

- A: α ≻ β ≻ γ
- B: β ≻ γ ≻ α
- C: γ ≻ α ≻ β

Let them vote between the three options. Two prefer α to β, and two prefer β to γ. We expect, then, that they will vote for α over γ. But in fact two prefer γ to α, so we have α ≻ β ≻ γ ≻ α. This is called a voting cycle or Condorcet cycle. Transitive individual preferences are aggregated via voting to an intransitive social preference in this case.

More generally consider a social choice rule \(\Phi\) as defined above satisfying the following properties:

1. Unrestricted Domain - it works for all possible N-tuples of preferences in \(\Pi^N\).
2. Pareto Principle: if all individual preferences prefer alternative α to alternative β then the social preference also prefers α to β.
3. Independence of Irrelevant Alternatives: the social preference between any two alternatives \{α, β\} depends only on individual's preferences over α and β and not on their preferences about other alternatives. Formally, for any pair of alternatives \(α, β \in \Pi\), and for any two preference n-tuples \(≽_i\) and \(≽'_i\), if \(≽_i\) and \(≽'_i\) agree on \{α, β\} then the social preference between α and β is the same for both preference N-tuples.
4. Non-Dictatorship: there is no individual such that \(\{α ≽_i β \Rightarrow α ≽ β\}\) where \(≽\) denotes the social preference.

**Theorem 1.** (Arrow’s Impossibility Theorem) There is no social choice rule satisfying the above four conditions.

**Proof.** First we prove a preliminary result, as follows. **Lemma:** For any alternative b, if every individual i ranks b either strictly best of all alternatives or strictly worst, then the social preference \(≽\) must rank b either strictly best or strictly worst.

Suppose the lemma were false. Then there would be some set of individual preferences and two other alternatives \(a, c\) such that \(a ≽ b ≥ c\) even though each person has b as her best or worst alternative.
Modify individual preferences as follows. For any person who has $b$ at the top of her list, move $c$ up to second on the list so that it is now strictly above $a$. And for anyone who has $b$ at the bottom, move $c$ to the top so that it is strictly above $a$. These alterations do not change anyone’s preferences between $b$ and $a$. So by IIA society still prefers $a$ to $b$. Similarly they don’t change anyone’s preferences between $c$ and $a$, so society still prefers $b$ to $c$. Now $a \succeq b$ and $b \succeq c$ implies $a \succeq c$ but everyone ranks $c$ above $a$ by construction so by Pareto $a \succ c$, a contradiction. So the lemma must be true.

Now let $b$ be any alternative and start with a set of individual preferences where everyone puts $b$ last. By Pareto the social preference must rank $b$ last too. Next change voter 1’s preference so that $b$ moves from last to first. By the lemma we know that the social preference must put $b$ either last or first. Change voter 2’s preference in the same way: again the social preference must put $b$ last or first. By the time we have changed all preferences in this way we know that the social preference must put $b$ first. Let the individual where the switch from bottom to top in the social preference occurs be $B$, that is the first preference list where the social preference puts $b$ top is the one where everyone up to and including $B$ puts $b$ top, and everyone after $B$ puts $B$ last.

Now call the list of preferences where everyone up to but not including $B$ puts $b$ first, profile 1, and the list where everyone up to and including $B$ puts $b$ first, profile 2.

Next we show that $B$ is a dictator: if the social preference is to satisfy IIA and transitivity then it always has to agree with $B$’s preference whatever anyone else thinks.

Consider any two alternatives $a, c$ which are different from $b$. We will show that if $a \succ_B c$ then $a \succ c$. We begin by considering a set of preferences in which $a \succ_B c$ and call this set of preferences profile 4. From this we construct profile 3 as follows. Move $a$ to the top of $B$’s preferences and $b$ to second. For individuals up to but not including $B$ move $b$ to the top. For individuals $B+1$ and onwards move $b$ to the bottom. This is profile 3.

Now in moving from profile 4 to profile 3 we did not change anyone’s ranking of $a$ versus $c$: for $B$ $a$ was above $c$ and we moved it to the top and for the others we only moved $b$. So by IIA the social preference between $a$ and $c$ has to be the same at profiles 3 and 4, and we will show that at profile 3 society prefers $a$ to $c$.

Recall that at profile 1 society puts $b$ last so that $a \succ b$. Everyone’s ranking of $a$ versus $b$ is the same in profiles 1 and 3. Hence by IIA $a \succ b$ at
At profile 2 society puts \( b \) first so \( b \succ c \) at profile 2. Everyone’s ranking of \( b \) versus \( c \) is the same in profiles 2 and everyone’s ranking of \( b \) versus \( c \) is the same in 2 and 3. So by IIA \( b \succ c \) at profile 3.

Since preferences are transitive we have that for profile 3 \( a \succ c \). Hence at profile 4 \( a \succ c \). So whenever \( a \succ_B c \), it follows that \( a \succ c \).

We have now shown that \( B \) is dictatorial over choices other than \( b \); next we show that she is dictatorial over choices involving \( b \) too. We need to show that if \( b \succ_B a \) then \( b \succ a \) and if \( a \succ_B b \) then \( a \succ b \).

To do this we pick an alternative different from \( a, b \) - say \( c \) - and repeat everything we have done so far: this will prove that someone is a dictator for any two alternatives that are not \( c \), that is she is a dictator when it comes to choices between \( a \) and \( b \). We need to show that this dictator is individual \( B \) again. We know that \( B \)'s preferences between \( a \) and \( b \) sometimes matter, because when we moved from profile 1 to 2 all that changed was \( B \)'s preference over \( b \), and in the social ranking \( b \) moved from bottom to top. Hence if the new dictator is not \( B \) we have a contradiction. Hence \( B \) is a dictator over any pair of alternative that excludes \( b \) and any that includes \( b \). This completes the proof. Any rule satisfying the first three conditions is dictatorial. \( \square \)

This is a very influential result - it has been taken as saying that perfect democracy is impossible. But there are social choice rules that satisfy points 2, 3 and 4 if we place some restrictions on the \( N \)-tuples of preferences admitted and drop the unrestricted domain condition. There are many results showing that there is a social choice rule satisfying 2, 3 and 4 if the preferences we consider are all in some way similar. The classic case of similarity is single-peakedness: choices are over a naturally ordered one dimensional variable (tax rate, budget deficit, ...) and each person has a most preferred value for this and ranks alternatives lower, the further they are from this most preferred value. In this case voting works, and the outcome will be the most preferred outcome of the median voter when voters are ranked by their most preferred outcomes. Here is a formal statement and proof.

**Definition 16.** Consider preferences over the real line \( \mathbb{R}^1 \). We say an ordering \( \succeq \) is single peaked if there exists an alternative \( m \) such that for \( y, z > m \), \( y \succeq z \) iff \( z > y \) and for \( y, z < m \), \( y \succeq z \) iff \( y > z \). A profile of preferences \( \succeq_i \) is single peaked if for every \( \succeq_i \) there exists such an \( m_i \).

Another related definition:
Definition 17. An agent \( k \) is the median agent for the preference profile \( \succeq_i, i \in I \), if \( N \{ i : m_i \geq m_k \} \geq \frac{I}{2} \) & \( N \{ i : m_i \leq m_k \} \geq \frac{I}{2} \).

Here \( N \{ X \} \) denotes the number of elements in the set \( \{ X \} \). Now we show that with single peaked preferences, the median voter’s preferred point is always chosen by a majority voting process.

**Proposition 19.** Consider a profile of single-peaked preferences with agent \( k \) the median agent. Then the peak of the median agent \( m_k \) cannot be defeated by any other alternative by majority voting.

**Proof.** Pick any \( y \) and assume that \( m_k > y \). Consider the set of agents \( S \subset I \) who have peaks greater than or equal to \( m_k \), \( \{ i \in I : m_i \geq m_k \} \). Then \( m_i \geq m_k > y \forall i \in S \). Hence \( m_k \succeq_i y \forall i \in S \). As agent \( k \) is a median agent, \( N \{ S \} \geq \frac{I}{2} \). So more than half the voters prefer \( m_k \) to \( y \). \( \square \)

7 Final comments on preferences and choice

7.1 Framing

We have set up a very formal and rational model of choice by consumers. Psychologists who study consumer behavior find this excessively rational. Choices are affected by **framing** - how an issue is presented. Here’s an example from Psychological Science, 21(1) pp 86-92 2010. A Dirty Word or a Dirty World? By David Hardisty, Eric Johnson and Elke Weber.

“Suppose you are purchasing a round trip flight from Los Angeles to New York City, and you are debating between two tickets, one of which includes a carbon tax [offset]. You are debating between the following two tickets, which are otherwise identical. Which would you choose? The ticket including the carbon tax [offset] costs $392.70 and the ticket without costs $385.” The results are striking and depend on whether the extra $7.70 – a mere 2% of the ticket costs – is described as a tax or an offset, and on the political leanings of the subjects. When the $7.70 is described as an offset, the proportions of democrats, independents and republicans agreeing to pay the extra amount are 56%, 49% and 53% respectively, roughly the same. But when the $7.70 was described as a tax, the results changed dramatically in the case of the independents and republicans: the numbers were now 50%, 28% and 13%. The acceptance rate almost halved for independents and dropped by 75% for republicans. Nothing changed except the frame of reference through which
people saw the issue and whether this triggered their hostility to taxes. There are many other cases of framing affecting a choice.

7.2 Endogenous Preferences

We have taken preferences as given, primitive data exogenous to the choice process. Perhaps preferences are in part at least formed by social and economic experience. Does this invalidate our approach? Not necessarily. In a more dynamic context we can model preferences that respond to consumption experience. For example, let \( c_t \) be a person’s consumption vector at time \( t \), and let preferences be such that they can be represented by the following utility function: \( U = \sum_t u(c_t) \delta^t \) with \( \delta \in [0,1] \) the discount factor. This is a standard model of preferences over time, and takes preferences in each period as exogenous and fixed. We can modify this to \( U = \sum_t u(c_t, z_t) \delta^t \) where \( z_t = \sum_{\tau=0}^{t-1} c_\tau \rho^{t-\tau} \). This defines a complete transitive preference ordering over consumption sequences, in which preferences today depend on past consumption experiences. (See paper by me and Harl Ryder, Rev Econ Stud 1973.)
Part II
Firms and Production Plans

Firms seek to maximize profits subject to various constraints so there is a lot of overlap with consumer theory. But there is no budget constraint, which makes matters easier.

We let \( y \in \mathbb{R}^N \) be a production plan, a list of inputs used and the outputs produced from them. **Sign convention:** inputs are negative and outputs positive, so \( \{2, 5, 7, -1, -9, 0\} \) means that 2, 5 and 7 units of goods one, two and three are produced and 1 and 9 units of four and five are used as inputs, with six neither consumed not produced. With this sign convention, \( p.y \) is profit: it is the cost of inputs subtracted from the value of outputs. So firms seek to maximize \( p.y \).

**Definition 18.** The production possibility set is the set of production plans that is feasible for the firm, denoted \( Y \subset \mathbb{R}^N \). Clearly \( y \in Y \).

The production possibility set is limited primarily by technology - you need iron ore to make steel, wheels to make a car, etc - but also by laws and regulations - you cannot dump toxic chemicals in the water, at least in the US - and by available resources - an oil company may run out of oil reserves.

We sometimes characterize the production possibility set in terms of a transformation function \( F : \mathbb{R}^N \rightarrow \mathbb{R} \), \( \{ Y = y \in \mathbb{R}^N : F(y) \leq 0 \} \). The boundary of the production set is known as the transformation frontier, \( \{ y : F(y) = 0 \} \).

The ratio

\[
\frac{\partial F}{\partial y_l} / \frac{\partial F}{\partial y_j} = MRT_{lj}
\]

is known as the marginal rate of transformation between \( l \) and \( j \). This number is the negative of the slope of the frontier.

If there is only one output \( y_1 \) we can write

\[
y_1 = f(y_2, ..., y_N)
\]

and the function \( f \) is known as the production function. For a given level of output \( y_1 \) we can look at the set of all inputs that can produce \( y_1 \):

\[
\{ y_2, ..., y_N : f(y_2, ..., y_N) \geq y_1 \}
\]
The boundary of this set is called the $y_1$-isoquant: \{\(y_2, \ldots, y_N : f(y_2, \ldots, y_N) = y_1\)\}. The slope of this isoquant is the marginal rate of technical substitution. A common example of a production function is the Cobb-Douglas:

\[ y = x_1^\alpha x_2^\beta \]

Here the marginal rate of technical substitution [MRTS] between \(x_1\) and \(x_2\) is

\[ \text{MRTS}_{1,2} = \frac{\alpha x_2}{\beta x_1} \]

### 7.3 Properties of the Production Set

1. \(Y\) is non-empty
2. \(Y\) is closed
3. No free lunch: \(Y \cap \mathbb{R}^N_+ \subset \{0\}\)
4. Possibility of inaction: \(0 \in Y\). Can conflict with sunk costs.
5. Free disposal: if \(y \in Y \& y' \leq y \Rightarrow y' \in Y\). Note that this can conflict with environmental regulations.
6. Irreversibility: \(y \in Y, y \neq 0 \Rightarrow -y \notin Y\)
7. Constant returns to scale: \(y \in Y \Rightarrow \alpha y \in Y \forall \alpha > 0\). Geometrically \(Y\) is a cone. Efficiency neither increases nor falls with scale.
8. Decreasing returns to scale: \(Y\) is strictly convex. Efficiency falls with scale.
9. Increasing returns to scale: \(y \in Y \Rightarrow \alpha y \in Y \forall \alpha > 1 \& \exists y \in Y : \beta y \notin Y, \beta \in (0, 1)\).
10. Additvity: \(y, y' \in Y \Rightarrow y + y' \in Y\)
11. Convexity: strict convexity implies diminishing returns to scale, convexity implies non-increasing returns to scale. (Non-increasing returns means \(\forall y \in Y, \beta y \in Y, \beta \in [0, 1]\).
7.4 Profit Maximization (PMP)

Problem is

\[ \max_y \{ p.y \}, \; y \in Y \]

or

\[ \max_y \{ p.y \}, \; F(y) \leq 0 \]

and given that the profit maximizing plan will generally be in the boundary we can work with

\[ \max_y \{ p.y \}, \; F(y) = 0 \]

**Definition 19.** Profit function \( \pi(p) = \max_y \{ p.y \}, \; y \in Y \)

**Definition 20.** Supply correspondence/function \( \{ y(p) : p.y = \pi(p) \} \)

Profit maximization involves maximizing the value of output net of the cost of inputs: we may write this problem in several ways.

\[ \max_x \{ p_y - p_x x \}, \; y = f(x), \; \max_y \{ p.y \}, \; F(y) = 0 \]

depending on whether we have a single output whose production can be expressed as a function of the inputs \( y = f(x) \) or many outputs so that we need to work with an implicit function \( F(y) = 0 \). For the general implicit function case the first order conditions are

\[ p_t = \lambda \frac{\partial F(y^*)}{\partial y_t} \]

where \( \lambda \) is a Lagrange multiplier. For the alternative case we can write the maximand as \( p_y f(x) - p_x x \) and the FOCs are

\[ p_y \frac{\partial f}{\partial x_t} \leq p_t, \quad \text{if } x_t > 0 \]

These FOCs are generally necessary conditions for profit maximization: they are also sufficient if the production possibility set is convex. From the FOCs it follows that the MRTS will equal the price ratio:

\[ \frac{\partial F/\partial x_t}{\partial F/\partial x_j} = \frac{p_t}{p_j} \]

and of course the FOC for the single output case states that the price of an input is the value of its marginal product.

Some facts about the profit function and the supply correspondence:
Proposition 20. (1) The profit function $\pi(p)$ is homogeneous of degree one.
(2) The supply correspondence $y(p)$ is homogeneous of degree zero.
(3) If $y(p)$ consists of a single point then $\pi(p)$ is differentiable at $p$ and
$\nabla \pi(p) = y(p)$ (Hotelling’s Lemma)
(4) If $y(p)$ is a function and differentiable at $p$, then $Dy(p) = D^2\pi(p)$ is
a symmetric and positive semi-definite matrix with $Dy(p)p = 0$.

Point (4) here is sometimes called the Law of Supply: it says that
quantities respond to a price change in the same direction as the price change
- if a price increases then the supply of that good increases too.

7.5 Cost Minimization (CMP)

If a firm is maximizing profits then there is no way of producing the same
output at lower cost, so it is also cost minimizing. Minimizing the cost of
producing output $y$ requires

$$\min_x \{p_x.x\}, \quad f(x) \geq y$$

The minimized value of this gives the cost function $c(p_x, y)$, the cost of
producing $y$ at prices $p_x$: $c(p_x, y) = \min_x p_x.x, \quad f(x) \geq y$. The FOCs for
this problem require that

$$\frac{f_j}{f_k} = \frac{p_j}{p_k}$$

so the MRTS equals the price ratio.

Proposition 21. (1) $c(p_x, y)$ is homogeneous of degree one in prices and
non-decreasing in output.
(2) $c(p_x, y)$ is a concave function of $p_x$
(3) If $f(x)$ is homogeneous of degree one (constant returns to scale) then
$c(p_x, q)$ is homogeneous of degree one in $y$.
(4) If $f(x)$ is concave then $c(p_x, y)$ is a convex function of $y$, which means
that marginal costs are non-decreasing in $y$.

Using the cost function we can restate the problem of maximizing profits
as:

$$\max_y \{py\} - c(p_x y)$$

and a FOC for this is clearly that

$$p \leq \frac{\partial c}{\partial y}, \quad \text{if } y > 0$$
which means that marginal cost should equal price.

7.5.1 Examples:

Cobb-Douglas Production Function

Consider the function $y = L^\alpha K^\beta$, $\alpha + \beta < 1$, which shows diminishing returns to scale. The PMP is

$$Max_{L,K} \{py\} - wL - rK, y = L^\alpha K^\beta \Leftrightarrow Max_{L,K} \{pL^\alpha K^\beta - wL - rK\}$$

and so the FOCs are

$$\alpha pL^{\alpha-1}K^\beta - w = 0$$
$$\beta pL^\alpha K^{\beta-1} - r = 0$$

which imply

$$\frac{\alpha K}{\beta L} = \frac{w}{r}$$

This yields

$$L^* = \left\{ \left( \frac{\alpha}{w} \right)^{1-\beta} \left( \frac{\beta}{r} \right)^\beta p \right\}^{\frac{1}{1-\alpha-\beta}}$$
$$K^* = \left\{ \left( \frac{\alpha}{w} \right)^\alpha \left( \frac{\beta}{r} \right)^{1-\alpha} p \right\}^{\frac{1}{1-\alpha-\beta}}$$

The supply function associated with these is

$$y^* = L^{*\alpha} K^{*\beta} = \left\{ \left( \frac{\alpha}{w} \right)^\alpha \left( \frac{\beta}{r} \right)^\beta p^{\alpha+\beta} \right\}^{\frac{1}{1-\alpha-\beta}}$$

Next look at the CMP:

$$Min_{L,K} \{wL + rK\}, L^\alpha K^\beta \geq y$$

The Lagrangian is

$$L = wL + rK + \lambda [y - L^\alpha K^\beta]$$

and the FOCs are

$$w - \lambda \frac{\alpha y}{L^*} = 0, \ r - \lambda \frac{\beta y}{K^*} = 0, \ y - L^{*\alpha} K^{*\beta} = 0$$
Dividing the first by the second and substituting into third yields factor demand functions given \( y \):

\[
L^* = \frac{\alpha}{k} \left( \frac{r}{w} \right)^{\frac{\beta}{\alpha + \beta}} y^{\frac{1}{\alpha + \beta}}, \quad K^* = \frac{r}{k} \left( \frac{\alpha}{\beta} \right)^{-\alpha/(\alpha + \beta)} y^{1/(\alpha + \beta)}
\]

where \( k = a^{(\alpha/\alpha + \beta)} \beta(\beta/\alpha + \beta) \).

We can now show that

\[
\lambda = \frac{1}{k} w^{\alpha/\alpha + \beta} r^{\beta/\alpha + \beta} y^{1/(\alpha + \beta)}
\]

The cost function is

\[
c(w, r, y) = wL^* + rK^* = \left( \frac{\alpha + \beta}{k} \right) w^{\alpha/\alpha + \beta} r^{\beta/\alpha + \beta} y^{1/(\alpha + \beta)}
\]

so the marginal cost is

\[
MC = \frac{\partial c}{\partial y} = \frac{1}{k} w^{\alpha/\alpha + \beta} r^{\beta/\alpha + \beta} y^{1/(\alpha + \beta)} = \lambda
\]

so the Lagrange multiplier from the CMP is the marginal cost.

We can compute the average cost

\[
AC = \frac{c(w, r, y)}{y} = Gy^{1-\alpha-\beta}/\alpha + \beta, \quad G = \left( \frac{\alpha + \beta}{k} \right) w^{\alpha/\alpha + \beta} r^{\beta/\alpha + \beta}
\]

and check whether this is increasing or decreasing with output:

\[
\frac{\partial AC}{\partial y} = \left( \frac{1 - \alpha - \beta}{\alpha + \beta} \right) G y^{\frac{1-\alpha-\beta}{\alpha + \beta}} < 0 \iff \alpha + \beta > 1
\]

so average costs are decreasing with output if and only if we have increasing returns to scale. With diminishing returns to scale average costs are rising with output level.

**Leontief technology**

The production function is

\[
f(x_1, x_2) = \min [ax_1, bx_2]
\]

As we know that the firm will not waste inputs, it must operate where \( y = ax_1 = bx_2 \). So the input demands are

\[
(x_1^*, x_2^*) = \left( \frac{y}{a}, \frac{y}{b} \right)
\]
and the cost function is
\[ c(w_1, w_2, y) = y \left( \frac{w_1}{a} + \frac{w_2}{b} \right) \]

**Example: Linear Technology**

Consider a linear production function
\[ f(x_1, x_2) = ax_1 + bx_2 \]
Goods are perfect substitutes and so the firm will use whichever is cheaper. Hence the input demands are
\[
(x^*_1, x^*_2) = \begin{cases} 
\left( \frac{y}{a}, 0 \right) & \text{if } \frac{w_1}{a} < \frac{w_2}{b} \\
\left( 0, \frac{y}{b} \right) & \text{if } \frac{w_1}{a} > \frac{w_2}{b} \\
\{ (x_1, x_2 : ax_1 + bx_2 = y, x_1, x_2 \geq 0) & \text{if } \frac{w_1}{a} = \frac{w_2}{b} \}
\end{cases}
\]

and the cost function is just
\[ c(w_1, w_2, y) = \min \left[ \frac{w_1}{a}, \frac{w_2}{b} \right] y \]
Figure 7.1

Figure 5.D.1
A strictly convex technology (strictly decreasing returns to scale).
(a) Production set.
(b) Cost function.
(c) Average cost, marginal cost, and supply.

Figure 5.D.2
A constant returns to scale technology.
(a) Production set.
(b) Cost function.
(c) Average cost, marginal cost, and supply.

Figure 5.D.3
A nonconvex technology.
(a) Production set.
(b) Cost function.
(c) Average cost, marginal cost, and supply.

Figure 5.D.4
Strictly convex variable costs with a nonsunk setup cost.
(a) Production set.
(b) Cost function.
(c) Average cost, marginal cost, and supply.

Figure 5.D.5
Constant returns variable costs with a nonsunk setup cost.
(a) Production set.
(b) Cost function.
(c) Average cost, marginal cost, and supply.
7.6 Aggregation of Firms

We noted that it is hard to aggregate consumers. It is much easier to aggregate firms. In this context a useful result is the following, which tells us that maximizing profits over the sum of all firms’ production possibility sets is equivalent to maximizing over each set individually and then adding up these maxima:

**Proposition 22.** Let $Y_i, i = 1, \ldots, I$ be production possibility sets and $Y = \sum_i Y_i$ be their sum. $p \in R^N$ is a price vector. Let $y^* = \arg\max_{y \in Y} \{p.y\}$ and $y_i^* = \arg\max_{y_i \in Y_i} \{p.y_i\}$. Then $y^* = \sum_i y_i^*$.

**Proof.** $\sum_i y_i^* \in \sum_i Y_i = Y$. Any $y \in Y$ can be written as $y = \sum_i y_i$, $y_i \in Y_i$. But $p.y_i^* \geq p.y_i \ \forall y_i \in Y_i, \ \forall i$. So $\sum_i p.y_i^* = p.\sum_i y_i^* \geq p.\sum_i y_i \ \forall y_i \in Y_i \Rightarrow p.\sum_i y_i^* \geq p.y \ \forall y \leq Y$. \qed

44
8 Second Problem Set

1. Show that the production function \( y = f(x) \) has constant returns to scale if and only if it is homogeneous of degree one.

2. For a Cobb-Douglas production function of two arguments with the sum of the exponents less than one, show that input demands are decreasing in their own prices.

3. For the same production function as 2 above, show that the cross price derivatives of input demands are equal, i.e. if \( K, L \) are the two inputs and \( w, r \) their prices then \( \frac{\partial K}{\partial w} = \frac{\partial L}{\partial r} \).

4. For the same production function as 2 above derive the supply function.

5. Suppose \( f(z) \) is a concave production function with \( L - 1 \) inputs \((z_1, z_2, ..., z_{L-1})\). Suppose \( \frac{\partial f}{\partial z_l} > 0 \forall l, z \geq 0 \) and that the matrix \( D^2 f(z) \) is negative definite for all \( z \). Use the firm’s first order condition and the implicit function theorem to prove the following statements: (A) An increase in the output price always increases the profit-maximizing level of output. (B) An increase in output price increases the demand for some input. (C) An increase in the price of an input reduced the demand for that input.
Part III
Choice under Uncertainty

Definition 21. A simple lottery \( L \) is a list of \( N \) exclusive and exhaustive outcomes \( 1, \ldots, N \) with associated probabilities \( (p_1, p_2, \ldots, p_N) \), \( \sum_n p_n = 1, p_n \in [0, 1] \), where \( p_n \) is the probability of outcome \( n \) occurring.

A simple lottery can be represented by a point in the \( N - 1 \)-dimensional simplex \( \Delta = \{ p \in R^N_+ : \sum_n p_n = 1 \} \).

Definition 22. Given \( K \) simple lotteries \( L_k = (p_{k1}, \ldots, p_{kN}) \), \( k = 1, \ldots, K \), and probabilities \( 1 \geq a_k \geq 0, \sum_k a_k = 1 \), the compound lottery \( (L_1, \ldots, L_K; a_1, \ldots a_K) \) is the risky alternative that yields the simple lottery \( L_k \) with probability \( a_k \).

For any such compound lottery we can calculate a corresponding reduced lottery as the simple lottery \( L = (p_1, \ldots, p_N) \) that generates the same ultimate distribution over outcomes. The probability of each outcome \( 1, \ldots, N \) is found by multiplying the probability of each lottery \( a_k \) by the probability \( p_{kn} \) that outcome \( n \) occurs in lottery \( k \), and then adding over lotteries \( k \). So the probability of outcome \( n \) in the reduced lottery is

\[
p_n = a_1 p_{n1}^1 + a_2 p_{n2}^2 + \ldots + a_K p_{nK}^K = \sum_k a_k p_{kn}^k
\]

9 Preferences over Lotteries

We assume the preferences over simple or compound lotteries depend only on the outcomes and their probabilities, and not in any way on the process reached to arrive at these outcomes and probabilities. So we take the set of alternatives to be the set of all simple lotteries \( L \) over the set of outcomes \( C \). People are assumed to have rational (complete, transitive) preferences \( \succeq \) over \( L \). Additional technical assumptions are:

Definition 23. The preference relation \( \succeq \) on the space of simple lotteries \( L \) is continuous if for any \( l, l', l'' \in L \) the sets

\[
\{ a \in [0, 1] : a l + (1 - a) l' \succeq l'' \} \subset [0, 1]
\]

and

\[
\{ a \in [0, 1] : l'' \succeq a l + (1 - a) l' \} \subset [0, 1]
\]
are closed.

So the set of combinations of \( l, l' \) that are at least as good as \( l'' \) is a closed set, as is the set that of combinations that are no better than \( l'' \). As in the deterministic case this is ruling out lexicographic preferences where the agent places all emphasis on the probability of one particular outcome - for example on the risk of death being zero. Next comes another assumption, a very crucial one.

**Definition 24.** The preference relation \( \succeq \) on the space of simple lotteries \( L \) satisfies the independence axiom if for all \( l, l', l'' \in L \) and \( a \in (0, 1) \) we have \( l \succeq l' \) iff \( al + (1 - a) l'' \succeq al' + (1 - a) l'' \).

In words, if we mix two lotteries with a third one, the the preference ordering of the two resulting mixtures does not depend on (is independent of) the particular third lottery used. \( l \) is preferred to \( l' \) iff \( l \) in any mixture with a third lottery \( l'' \) is preferred to \( l' \) in the same mixture with \( l'' \). Here’s an example of what this axiom means.

Suppose \( l \succeq l' \) and \( a = 0.5 \). \( 0.5l + 0.5l'' \) is the lottery resulting from a coin toss between \( l, l'' \), say heads giving \( l \) and tails \( l'' \). Likewise \( 0.5l' + 0.5l'' \) is the lottery generated by a coin toss between \( l', l'' \), again heads giving \( l' \). Conditional on heads \( 0.5l + 0.5l'' \) is at least as good as \( 0.5l' + 0.5l'' \), and conditional on tails they give the same outcome. So it is reasonable that \( 0.5l + 0.5l'' \succeq 0.5l' + 0.5l'' \), which is what the axiom implies.

**Definition 25.** The utility function \( U : L \to R \) has an expected utility form if there is an assignment of numbers \( u_1, ..., u_N \) to the \( N \) outcomes such that for every simple lottery \( l = (p_1, ..., p_N) \in L \) we have

\[
U(l) = \sum_n u_n p_n
\]

A utility function with the expected utility form is called a von Neumann-Morgenstern (vN-M) expected utility function.

Note that if we let \( l^n \) denote the degenerate lottery yielding outcome \( n \) with probability one then \( U(l^n) = u_n \). The expression \( U(l) = \sum_n u_n p_n \) is linear in probabilities, suggesting
Proposition 23. A utility function $U : L \rightarrow R$ has an expected utility form if and only if it is linear in probabilities, that is

$$U\left(\sum_{k=1}^{K} a_k l_k\right) = \sum_{k=1}^{K} a_k U\left(l_k\right)$$

for any $K$ lotteries $l_k$ and probabilities $a_1, ..., a_K \geq 0, \sum_k a_k = 1$. [In words the utility of a mixture of simple lotteries equals the expectation of the utilities of the individual simple lotteries.]

Proof. Suppose $U$ has the linearity property. We can write any lottery $l = (p_1, ..., p_N)$ as a combination of the degenerate lotteries $l^1, \ldots, l^N$: $l = \sum_n p_n l^m$. Then $U(l) = U(\sum_n p_n l^m) = \sum_n p_n U(l^m) = \sum_n p_n u_n$. So $U$ has the expected utility form.

Now suppose $U$ has the expected utility form, and consider a compound lottery $(l_1, ..., l_K; a_1, ..., a_K)$ where $l_k = (p_{1}^{k}, \ldots, p_{N}^{k})$. Its reduced lottery is $l' = \sum_k a_k l_k$. Then we have, remembering that the probability of outcome $n$ in the reduced lottery is

$$p_n = a_1 p_n^1 + a_2 p_n^2 + \ldots + a_K p_n^K = \sum_k a_k p_n^k$$

$$U\left(\sum_k a_k l_k\right) = \sum_n u_n \left[\sum_k a_k p_n^k\right] = \sum_k a_k \left[\sum_n u_n p_n^k\right] = \sum_k a_k U\left(l_k\right)$$

which completes the proof. \qed

The utility functions we discussed in the section on consumer preferences were unique only up to a monotone transformation, that is, an order-preserving transformation. They are ordinal functions. That is not true of v.N-M utility functions: there are unique up to a linear transformation, and are cardinal functions.

Proposition 24. Suppose that $U : L \rightarrow R$ is a vN-M expected utility for the preference relation $\succeq$ on $L$. Then $\hat{U}$ is another vN-M function for $\succeq$ if and only if there exist scalars $\beta > 0, \gamma$ such that $\hat{U}\left(l\right) = \beta U\left(l\right) + \gamma \forall l \in L$. 

48
Proof. Choose two lotteries \( \bar{l}, l \) such that \( \bar{l} \succeq l \succeq l \) \( \forall l \in L \). If \( \bar{l} \sim l \) all lotteries are indifferent and the result is trivial. So we assume \( \bar{l} \succeq l \).

Note that if \( U \) is a vN-M function and \( \hat{U} = \beta U + \gamma \) then

\[
\hat{U} \left( \sum_k a_k l_k \right) = \beta U \left( \sum_k a_k l_k \right) + \gamma = \beta \left[ \sum_k a_k U(l_k) \right] + \gamma
\]

\[
= \sum_k a_k \left[ \beta U(l_k) + \gamma \right] = \sum_k a_k \hat{U}(l_k)
\]

so \( \hat{U} \) has the expected utility form.

For the reverse proof we need to show that if both \( U, \hat{U} \) have the expected utility form then there exist scalars \( \beta > 0, \gamma \) such that \( \hat{U} = \beta U + \gamma \). Consider any lottery \( l \) and define \( \lambda_l \) by

\[
U(l) = \lambda_l U(\bar{l}) + (1 - \lambda_l) U(l) = U(\lambda_l \bar{l} + (1 - \lambda_l) l)
\]

Hence

\[
\lambda_l = \frac{U(l) - U(\bar{l})}{U(l) - U(l)}
\]

Since \( \lambda_l U(\bar{l}) + (1 - \lambda_l) U(l) = U(\lambda_l \bar{l} + (1 - \lambda_l) l) \) and \( U \) represents the preference \( \succeq \) it follows that \( l \sim \lambda_l \bar{l} + (1 - \lambda_l) l \). In this case since \( \hat{U} \) is also linear and represents the same preferences

\[
\hat{U}(l) = \hat{U}(\lambda_l \bar{l} + (1 - \lambda_l) l) = \lambda_l \hat{U}(\bar{l}) + (1 - \lambda_l) \hat{U}(l) = \lambda_l \left[ \hat{U}(\bar{l}) - \hat{U}(l) \right] + \hat{U}(l)
\]

Substituting and rearranging terms shows that \( \hat{U}(l) = \beta U(l) + \gamma \) where

\[
\beta = \frac{\hat{U}(\bar{l}) - \hat{U}(l)}{U(\bar{l}) - U(l)}, \quad \gamma = \hat{U}(l) - U(l) \beta
\]

\( \Box \)

A result of this proposition is that for vN-M utilities, utility differences have meaning - which was not true for ordinal utilities. So if there are four outcomes it is meaningful to say “the utility difference between outcomes 1 and 2 is greater than that between outcomes 3 and 4,” which corresponds to
the statement \( u_1 - u_2 > u_3 - u_4 \). This in turn is equivalent to \( 0.5u_1 + 0.5u_4 > 0.5u_2 + 0.5u_3 \), which in turn means that the lottery \((0.5, 0, 0, 0.5)\) is preferred to \((0, 0.5, 0.5, 0)\).

A final point to note before we get to the most important result of this section is that the independence axiom implies that indifference curves are straight lines in prospect space. To see this note that an indifference curve is a straight line if \( l_1 \sim l_2 \Rightarrow al_1 + (1 - a)l_2 \sim l_1 \sim l_2 \forall a \in [0, 1] \). Suppose this is not the case and in fact \( 0.5l_1 + 0.5l_2 > l_2 \). This is the same as \( 0.5l_1 + 0.5l_2 > 0.5l_2 + 0.5l_2 \). But since \( l_1 \sim l_2 \), by the independence axiom, we must have \( 0.5l_1 + 0.5l_2 \sim 0.5l_2 + 0.5l_2 \). So indifference curves are linear: we can also show that they are parallel straight lines: it is easy to construct a contradiction if they are not.

**Proposition 25.** [Expected utility theorem.] Suppose that the rational preference relation \( \succeq \) on \( L \) satisfies the continuity and independence axioms. Then \( \succeq \) admits a representation in the expected utility form, that is we can assign numbers \( u_n \) to each outcome \( 1, \ldots, N \) in such a manner that for any two lotteries \( l = (p_1, \ldots, p_N), l' = (p'_1, \ldots, p'_N) \) we have

\[
l \succeq l' \iff \sum_n u_n p_n \geq \sum_n u_n p'_n
\]

**9.1 Paradoxes**

Lots of people feel uneasy about the independence axiom and there are various paradoxes that seem to violate it. One of the best known is the Allais Paradox. Number of outcomes is 3, \( N = 3 \).

<table>
<thead>
<tr>
<th>First prize</th>
<th>Second prize</th>
<th>Third prize</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2,500,000</td>
<td>$500,000</td>
<td>$0</td>
</tr>
</tbody>
</table>

First choice is between two lotteries \( l_1, l'_1 \):

\[
l_1 = 0, 1, 0 : \ l'_1 = 0.10, 0.89, 0.01
\]

Second is between \( l_2, l'_2 \)

\[
l_2 = 0, 0.11, 0.89 : \ l'_2 = 0.10, 0, 0.90
\]
It is common in experimental situations for people to rank $l_1 > l'_1 : l'_2 > l_2$.

The first choice implies the certainty of $500,000 is preferred to a lottery offering a 10% chance of five times as much together with a small risk of getting zero. The second implies that a 10% chance of getting $2,500,000 beats an 11% chance of $500,000.

These choices are inconsistent with a vN-M utility function. To see this denote by $u_{25}, u_5, u_0$ the utilities of the three outcomes. Then $l_1 > l'_1$ implies

$$u_5 > 0.1u_{25} + 0.89u_5 + 0.01u_0$$

Adding $0.89u_0 - 0.89u_5$ to both sides gives

$$0.11u_5 + 0.89u_0 > 0.1u_{25} + 0.9u_0$$

and therefore $l_2 > l'_2$. Several possible reactions to this.

1. People will change their choices if the inconsistency with the underlying axioms is shown to them.

2. This paradox is not very relevant because it involves probabilities near zero and one, and vastly different outcomes.

3. Regret is relevant: we prefer $l_1$ to $l'_1$ because we will always regret not getting $500,000 when we could have got that with certainty if we choose $l'_1$ and get zero. In the second case there is a good chance of getting zero anyway.

Another interesting paradox is the Ellsberg Paradox, as follows. There is an urn with 90 balls in it, 30 red and the rest black and yellow. So the probabilities of black and yellow are unknown - uncertainty or ambiguity rather than risk. People are given the following choices between pairs of lotteries:

| Gamble A: $10 of you draw red. |
| Gamble B: $10 if you draw black. |
| Gamble C: $10 for red or yellow |
| Gamble D: $10 for black or yellow |

Generally we find that people prefer A to B and also D to C. This also implies a violation of the independence axiom, as follows:

$$A > B \Rightarrow Ru_{10} + (1 - R) u_0 > Bu_{10} + (1 - B) u_0 \Rightarrow R > B$$
\[ D > C \Rightarrow Bu_{10} + Yu_{10} + Ru_0 > Ru_{10} + Yu_{10} + Bu_0 \Rightarrow B > R \]

So there is no set of probabilities and utilities that supports this choice as expected utility maximization. Choices under ambiguity seem to violate the independence axiom.

10 Risk Aversion

Consider a lottery over monetary amounts, non-negative numbers, with probabilities given by the density function \( f(t), t \in [0, \infty] \). The cumulative distribution function is \( F(x) = \int_0^x f(t) \, dt \). Note that the final distribution for a compound lottery is a weighted average of the distributions of each of the component lotteries: if \( l_1, \ldots, l_K : a_1, \ldots, a_K \) is a compound lottery then the cumulative distribution is \( F(x) = \sum_k a_k F_k(x) \).

We now take the set of all lotteries to be the set of all cumulative distributions over non-negative amounts of money. We can apply the vN-M theorem to show that there is a utility function of the form

\[
U(F) = \int u(x) \, dF(x)
\]

Note that \( U \) is defined on lotteries and \( u \) is defined on amounts of money. \( U \) is generally called the VN-M utility and \( u \) the Bernoulli utility. The axioms of expected utility theory place restrictions on \( U \) but not on \( u \), which could be any increasing continuous function.

Note: it is sometimes argued that \( u \) should be bounded above, because of the St Petersburg Paradox. Assume that \( u \) is unbounded and let \( x_m \) be an amount of money such that \( u(x_m) > 2^m \). Consider the following lottery:

Toss a coin repeatedly until heads comes up. If this happens on the \( m \)-th toss the payoff is \( x_m \). The expected utility from this lottery is

\[
\sum_{1}^{\infty} u(x_m) \left( \frac{1}{2} \right)^m > \sum_{1}^{\infty} 2^m \left( \frac{1}{2} \right)^m = +\infty
\]

so you ought in principle to be willing to pay any amount to play this lottery. Clearly most people are not, so this is an argument for \( u \) being bounded above.

**Definition 26.** A decision maker is risk averse if for any lottery \( F(.) \) the degenerate lottery that yields the expected amount \( \int xdF(x) \) with certainty
is at least as good as \( F \). If for all \( F \) the decision-maker is indifferent between these two lotteries we say he is risk-neutral, and we say he is strictly risk averse if he is indifferent only when the two lotteries are the same.

So a decision maker is risk averse if and only if
\[
\int u(x) \, dF(x) \leq u \left( \int x \, dF(x) \right) \quad \forall F
\]

In words, the expected utility of the outcome does not exceed the utility of the expected outcome. This is called Jensen’s Inequality, and is the inequality used to define a concave function, so risk aversion is equivalent to the function \( u \) being concave - diminishing marginal utility of income or wealth.

**Definition 27.** The certainty equivalent of a lottery \( F \), denoted \( c(F,u) \), is the amount of money that the individual regards as indifferent to the gamble represented by the lottery:

\[
u(c(F,u)) = \int u(x) \, dF(x)
\]

If the decision-maker is an expected utility maximizer with Bernoulli utility function \( u \) on amounts of money then:

**Proposition 26.** The following are equivalent:
1. The decision-maker is risk-averse
2. \( u(\cdot) \) is concave
3. \( c(F,u) \leq \int x \, dF(x) \) for all \( F \)

So risk-aversion, concavity and the certainty equivalent being less than the expectation are all equivalent.

### 10.1 Risk Management

We will look at insurance and portfolio choice.

**Insurance:** a strictly risk-averse person with initial wealth \( W \) runs a risk of losing \( D \) with probability \( \pi \). A unit of insurance costs \( q \) and pays \$1 if the loss occurs. So if \( a \) units of insurance are purchased then her wealth is \( W - aq \) if there is no loss and \( W - aq - D + a \) if the loss occurs. So expected wealth is

\[
(W - aq)(1 - \pi) + (W - aq - D + a) \pi = W - \pi D + a(\pi - q)
\]
Utility maximization requires

\[ \max_a (1 - \pi) u(W - aq) + \pi u(W - aq - D + a) \]

FOC is

\[-q(1 - \pi) u'(W - a^*q) + \pi (1 - q) u'(W - D + a^*(1 - q)) \leq 0, \quad \text{if } a^* > 0 \]

Assume the price of insurance is **actuarially fair**, that is it is equal to the expected cost of the insurance. This means \( q = \pi \). Then the FOC requires

\[ u'(W - D + a^*(1 - \pi)) - u'(W - a^*\pi) \leq 0, \quad \text{if } a^* > 0 \]

Since \( u'(W - D) > u'(W) \) it must be the case that \( a^* > 0 \) and so

\[ u'(W - D + a^*(1 - \pi)) = u'(W - a^*\pi) \]

Because \( u' \) is strictly decreasing in its argument this means that

\[ W - D + a^*(1 - \pi) = W - a^*\pi \]

or equivalently

\[ a^* = D \]

which means that the agent insures fully. So if the insurance is actuarially fair the risk-averse agents insures fully. Her wealth is then \( W - \pi D \) whether the loss occurs or not.

**Demand for a risky asset**: There are two assets, a safe asset with a return of 1 per dollar invested, and a risky one with a random return of \( z \) per dollar invested. \( z \) has a distribution function \( F(z) \) which we assume satisfies \( \int z dF(z) > 1 \), so that its mean return exceeds that of the safe asset.

Wealth \( W \) can be invested in any way between the two assets, with \( a, b \) the amounts invested in the risky and safe assets respectively, with \( a + b = W \). For any realization of \( z \) the portfolio pays \( az + b \). The choice problem is

\[ \max_{a,b} \int u(az + b) dF(z) = \int u(W + a(z - 1)) dF(z), \quad 0 \leq a \leq W \]

The solution \( a^* \) must satisfy

\[ \phi(a^*) = \int u'(W + a^*[z - 1]) (z - 1) dF(z) \leq 0 \quad \text{if } a^* < W, \geq 0 \quad \text{if } a^* > 0 \]

Note that \( \int zdF(z) > 1 \Rightarrow \phi(0) > 0 \). So \( a^* = 0 \) cannot satisfy this equation and the optimal portfolio has \( a^* > 0 \). Conclusion: if a risk is actuarially favorable then a risk averter will always accept at least a small amount of it.
Definition 28. Given a Bernoulli utility function $u$ for money, the Arrow-Pratt coefficient of absolute risk aversion at $x$ is defined as $r_A(x) = -u''(x)/u'(x)$. Note the negative sign in front: for concave functions this is always non-negative.

We know that risk neutrality is equivalent to linearity, and that risk aversion seems to increase with the curvature of $u$. The utility function can be recovered from $r_A$ by integrating twice, up to two integration constants. The integration constants don’t matter as $u$ is unique only up to two constants anyway.

Example: Consider the utility function $u(x) = -e^{-ax}$ for $a > 0$. Then we have $u'(x) = ae^{-ax}$ and $u''(x) = -a^2e^{-ax}$, so that $r_A(x, u) = a \forall x$.

Given two utility functions $u_1(x), u_2(x)$, when can we say that one is more risk averse than the other?

Proposition 27. The following are all equivalent:

1. $r_A(x, u_2) \geq r_A(x, u_1) \forall x$

2. There exists an increasing concave function $\psi(.)$ such that $u_2(x) = \psi(u_1(x)) \forall x$: that is $u_2$ is a concave transformation of $u_1$.

3. $c(F, u_2) \leq c(F, u_1) \forall F(.)$

4. Whenever $u_2(.)$ finds a lottery $F(.)$ at least as good as a riskless outcome $\bar{x}$, then $u_1(.)$ also finds $F(.)$ at least as good as $\bar{x}$. Or $\int u_2(x) dF(x) \geq u_2(\bar{x}) \Rightarrow \int u_1(x) dF(x) \geq u_1(\bar{x}) \forall F(.) , \bar{x}$.

Proof. Show that 1 and 2 are equivalent. Note that for some increasing function $\psi$ we always have $u_2(x) = \psi(u_1(x))$ because the two represent the same (increasing) ordering on $R^1$. Differentiating

$$u'_2(x) = \psi'(u_1(x))u'_1(x)$$

and again’

$$u''_2(x) = \psi'(u_1(x))u''_1(x) + \psi''(u_1(x))(u'_1(x))^2$$

Dividing both sides of $u''_2$ by $u'_2$ and using the first line we get

$$r_A(x, u_2) = r_A(x, u_1) - \frac{\psi''(u_1(x))}{\psi'(u_1(x))}u'_1(x)$$

From this we note that

$$r_A(x, u_2) \geq r_A(x, u_1) \text{ iff } \psi''(u_1) \leq 0$$

□
More-risk-averse-than is a transitive but incomplete ordering. Many Bernoulli utility functions will be incomparable, that is we will have \( r_A(x, u_1) > r_A(x, u_2) \) at some \( x \) and also \( r_A(x', u_1) < r_A(x', u_2) \) for some \( x' \neq x \).

**Example**

Next we will consider the portfolio choices of two risk-averse individuals and show that the more risk-averse will always invest less in the risk asset. As before, there are two assets, a safe asset with a return of \( l \) per dollar invested, and a risky one with a random return of \( z \) per dollar invested. \( z \) has a distribution function \( F(z) \) which we assume satisfies \( \int z dF(z) > 1 \), so that its mean return exceeds that of the safe asset.

Wealth \( W \) can be invested in any way between the two assets, with \( a_i, b_i \) the amounts invested in the risky and safe assets respectively, with \( a_i + b_i = W_i \). For any realization of \( z \) the portfolio pays \( a_iz + b_i \). The choice problem is

\[
Max_{a_i, b_i} \int u_i(a_iz + b_i) dF(z) = \int u_i(W_i + a_i(z - 1)) dF(z), \ 0 \leq a_i \leq W_i
\]

For interior solutions the solutions \( a_i^* \) must satisfy

\[
\phi_i(a_i^*) = \int u_i'(W_i + a_i^*[z - 1]) (z - 1) dF(z) = 0
\]

The concavity of \( u_2 \) implies that \( \phi_2 \) is decreasing, so if we show that \( \phi_2(a_1^*) < 0 \), it follows that \( a_2^* < a_1^* \), which is what we want to show.

Now \( u_2(x) = \phi(u_1(x)) \) where \( \phi \) is a concave function. Hence

\[
\phi_2(a_1^*) = \int (z - 1) \phi'(u_1(W_1 + a_1^*[z - 1])) u_1'(W_1 + a_1^*[z - 1]) dF(z) < 0
\]

The final inequality follows from the first order condition recalling that in this case we have the FOC multiplied by \( \phi' \) a positive decreasing function of \( z \).

**Definition 29.** The Bernoulli utility function \( u(\cdot) \) for money exhibits decreasing absolute risk aversion if \( r_A(x, u) \) is a decreasing function of \( x \).

People whose preferences show decreasing absolute risk aversion take more risk as they become richer.
Proposition 28. The following properties are equivalent:

1. The Bernoulli utility function \( u \) exhibits decreasing absolute risk aversion

2. Whenever \( x_2 < x_1 \), \( u_2 (z) = u(x_2 + z) \) is a concave transformation of \( u_1 (z) = u(x_1 + z) \)

3. For any risk \( F(z) \) the certainty equivalent of the lottery formed by adding risk \( z \) to wealth level \( x \), given by the amount \( c_x \) at which \( u(c_x) = \int u(x + z) dF(z) \), is such that \( x - c_x \) is decreasing in \( x \). So the higher is \( x \), the less the person is willing to pay to remove the risk.

4. For any \( F(x) \), if \( \int u(x_2 + z) dF(z) \geq u(x_2) \) and \( x_2 < x_1 \) then \( \int u(x_1 + z) dF(z) \geq u(x_1) \).

Next we look at another concept of risk aversion, relative rather than absolute. This is a measure of aversion to proportional fluctuations in wealth, rather than absolute fluctuations.

Definition 30. The coefficient or index of relative risk aversion (RRA) is

\[ r_R(x, u) = -xu''/u'. \]

Decreasing relative risk aversion \( [r_R \text{ decreasing with } x] \) means that a person becomes less averse to a given proportional risk as her income rises.

Examples of constant IRRA utility functions:

\[ u(C) = \log C, \quad u' = 1/C, \quad u'' = -1/C^2, \Rightarrow \text{IRRA} = 1 \]

\[ u(C) = C^{1-\eta}/(1-\eta), \quad u' = C^{-\eta}, \quad u'' = -\eta C^{-\eta-1}, \Rightarrow \text{IRRA} = \eta. \]

If \( \eta < 1 \) then \( u(C) > 0 \) and \( u \) is unbounded: if \( \eta > 1 \) then \( u < 0 \) and \( u \) is bounded.

Proposition 29. The following are equivalent for a Bernoulli utility function \( u(\cdot) \):

1. \( r_R(x, u) \) is decreasing in \( x \)

2. Whenever \( x_2 < x_1 \), \( \tilde{u}_2(t) = u(tx_2) \) is a concave transform of \( \tilde{u}_1(t) = u(tx_1) \)

3. Given any risk \( F(t) \) on \( t > 0 \), the certainty equivalent \( \tilde{c}_x = \int u(tx) dF(t) \) satisfies \( x/\tilde{c}_x \) is decreasing in \( x \).

Proof. We will show that 1. implies 3. Pick a distribution \( F(t) \) over \( t \), and for any \( x \) define \( u_x(t) = u(tx) \). Let \( c(x) \) be the usual certainty equivalent from definition 27: \( u_x(c(x)) = \int u_x(t) dF(t) \). Note that as \( u_x(t) = xu'(tx) \)

\[ -\frac{u_x''(t)}{u_x'(t)} = -\frac{1}{tx} \frac{u''(tx)}{u'(tx)} \]
for any $x$. Hence if 1. holds then $u_{x'}$ is less risk averse than $u_x$ whenever $x' > x$. It follows from proposition 27 that $c(x') > c(x)$ and so $c$ is increasing. By the definition of $u_x$, $u_x(c(x)) = u(xc(x))$. In addition

$$u_x(c(x)) = \int u_x(t) \, dF(t) = \int u(tx) \, dF(t) = u(\tilde{c}_x)$$

Hence $\tilde{c}_x/x = c(x)$ and so $x/\tilde{c}_x$ is decreasing, completing the proof.

### 10.1.1 Mean-Variance

Next a simple illustration of the role that the index of RRA can play. Let $y$ be a random variable distributed as $F(y)$ with mean $y^*$. The expected utility associated with this is $Eu = \int u(y) \, dF(y)$ and define $x$ so that

$$u(y^* - x) = Eu = \int u(y) \, dF(y)$$

Here $x$ is the cost of risk bearing, the difference between the certainty equivalent and mean $y^*$ of the prospect $F$. Clearly

$$u(y^* - x) - u(y^*) = \int [u(y) - u(y^*)] \, dF(y) = \int [u'(y^*)(y - y^*) + \frac{u''(y^*)}{2}(y - y^*)^2] \, dF(y)$$

so that

$$u(y^* - x) - u(y^*) = \int \left[ \frac{u''(y^*)}{2}(y - y^*)^2 \right] \, dF(y) = \frac{u''(y^*) \sigma^2}{2}$$

and we can rewrite the LHS to give

$$-u'(y^*)x = \frac{u''(y^*) \sigma^2}{2}$$

$\Rightarrow x = -\frac{1}{2} \frac{y^* u''(y^*) \sigma^2}{u'(y^*) y^*} = \frac{1}{2} \frac{\sigma^2}{y^*}$$

So the cost of risk bearing is one half of the variance of the risk over the mean outcome times the IRRA. It is also the index of absolute risk aversion times half the variance.

We can write

$$Eu = u(y^* - x) = u(y^*) + \frac{1}{2} u''(y^*) \sigma^2$$
which is a function of the mean outcome \( y^* \) and its variance \( \sigma^2 \), and from this we can use the implicit function theorem to get

\[
\frac{\partial y^*}{\partial \sigma} = -\frac{u'' \sigma}{u' + 0.5u'''}
\]

Assuming \( u''' \approx 0 \) so the utility is almost quadratic we have

\[
\frac{\partial y^*}{\partial \sigma} = -\frac{u'' \sigma}{u'} > 0
\]

as the slope of an indifference curve in \( y^* - \sigma \) space. So this is linear if the index of absolute risk aversion (IARA) is constant and defines convex preferred-or-indifferent sets if the IARA is increasing.

10.2 Comparison of payoffs in terms of return and risk

In comparing risky choices, we can ask two different questions: is one more rewarding than the other, in terms of offering better outcomes, and is one more risky than the other?

First we formalize the idea that distribution \( F \) yields unambiguously higher returns than distribution \( G \). We assume distributions satisfy \( F(0) = 0 \) and \( F(x) = 1 \) for some \( x \). Two possible approaches: one to ask whether every expected utility maximizer whose utility is increasing in income will prefer one to the other, and the second is to ask if for every amount of money \( x \) the probability of getting at least \( x \) is greater under one than under the other.

Both approaches lead to the same concept.

**Definition 31.** The distribution \( F \) first order stochastically dominates distribution \( G \) if, for every non-decreasing function \( u : R \to R \),

\[
\int u(x) \, dF(x) \geq \int u(x) \, dG(x)
\]

**Proposition 30.** The distribution \( F \) first order stochastically dominates distribution \( G \) if and only if \( F(x) \leq G(x) \) \( \forall x \)

**Proof.** Note first that if \( f(x), g(x) \) are the pdfs of \( F(x), G(x) \) and \( [a, b] \) contains the supports of both distributions then integrating by parts

\[
\int_a^b u(x) f(x) \, dx = [u(x) F(x)]_a^b - \int_a^b u'(x) F(x) \, dx
\]
which reduces to
\[ u(b) - u(a) = \int_a^b u'(x) F(x) \, dx \]

So comparing the expected utility under the two distributions gives
\[
\int_a^b u(x) f(x) \, dx - \int_a^b u(x) g(x) \, dx = -\int_a^b u'(x) F(x) \, dx + \int_a^b u'(x) G(x) \, dx
\]
\[ = \int_a^b u'(x) [G(x) - F(x)] \, dx
\]

We want this to be positive for all increasing functions \( u : u'(x) > 0 \forall x \in [a, b] \). Clearly this is true if \( G(x) > F(x) \forall x \in (a, b) \). So we have shown that \( G(x) \geq F(x) \) implies that \( F \) first order stochastically dominates \( G \).

Now the reverse: \( G(x) \geq F(x) \iff F \text{ FOSD } G \). Suppose to the contrary that \( \exists x' : F(x') > G(x') \). Then we can construct a function \( u \) for which \( u' \) is very large where \( F > G \) and very small elsewhere. This makes the integral on the RHS negative, so that \( F \) does not dominate \( G \).

Next we turn to a discussion of when one distribution is more risky than another, and of second order stochastic dominance. We compare only distributions with the same mean.

**Definition 32.** For any two distributions \( F \) and \( G \) with the same mean, \( F \) second order stochastically dominates \( G \) (or is less risky than \( G \)) if for every non-decreasing concave function \( u : R_+^N \rightarrow R_+^N \) we have
\[
\int u(x) \, dF(x) \geq \int u(x) \, dG(x)
\]

Next we discuss an alternative way of characterizing second order stochastic dominance, using the idea of a mean-preserving spread.

**Definition 33.** Distribution \( G \) is a mean-preserving spread of distribution \( F \) if \( G \) is the reduction of a compound lottery made up of the distribution \( F \) with an additional lottery so that when \( F \) selects \( x \) the final outcome is \( x + z \) where \( z \) is a random variable whose mean is zero.

**Proposition 31.** Consider two distributions \( F \) and \( G \) with the same mean. Then the following statements are equivalent.
1. \( F(\cdot) \) second order stochastically dominates \( G(\cdot) \)
2. \( G(\cdot) \) is a mean-preserving spread of \( F(\cdot) \)
3. \( \int_0^x G(t) \, dt \geq \int_0^x F(t) \, dt \forall x \)
10.3 A Geometric Approach to Insurance

Figure 10.1:

Figure 10.1 gives a geometric way of thinking about insurance. There are two states, 1 & 2. It is not certain which will occur, and their respective probabilities are $p_1, p_2$. The consumer’s initial endowment is at the point $z_1$ giving $z_{11}$ in state 1 and $z_{12}$ in state 2. The 45 degree line shows situations where income is the same in each state, and these are therefore fully-insured positions. The consumer’s expected utility is given by

$$u(z_{11})p_1 + u(z_{12})p_2$$

and the slope of an indifference curve is therefore

$$-\frac{p_1u'(z_{11})}{p_2u'(z_{12})}$$
On the 45 degree line, \( z_{11} = z_{12} \) so this slope is just \(-p_1/p_2\), the ratio of the probabilities.

Now consider the move from the initial position \( z_1 \) to the fully insured position \( z_2 \). This involves selling \( z_1 - z_0 \) of income in state 1 and buying \( z_2 - z_0 \) of income in state 2. This transaction will move the consumer to a fully insured position. What is the expected value of this transaction? The probability of state 1 is \( p_1 \) so the probability of giving up \( z_1 - z_0 \) is \( p_1 \), and the probability of state 2 and so of acquiring \( z_2 - z_0 \) is \( p_2 \). So the expected value of this transaction is \(-p_1 (z_1 - z_0) + p_2 (z_2 - z_0)\) which is zero if \( \frac{p_1}{p_2} = \frac{z_2 - z_0}{z_1 - z_0} \), so the transaction is actuarially fair if the slope of the budget line, which is the right hand side here, equals the probability ratio. As the slope of an indifference curve is always equal to the probability ratio on the 45 degree line, the offer of actuarially fair insurance will always be accepted and lead to full insurance.

Note that in this context convexity of the preferred-or-indifferent sets is equivalent to risk aversion: it implies a preference for moving towards the 45 degree line.

## 11 Discount Rates and the Elasticity of Marginal Utility

We have discussed the index of relative risk aversion, \(-xu''(x)/u'(x)\). This parameter is also important in other areas of economics, where it is known as the elasticity of the marginal utility of consumption and generally denoted \( \eta(c) \). To see why it is called this note that

\[
\frac{du'(c)}{dc} \frac{c}{u'(c)} = \frac{cu''}{u'}
\]

Note also that the proportional rate of change of the present value of marginal utility \( u'(c_t)e^{-\delta t} \) (where \( c_t \) is a function of time \( t \)) is given by

\[
dln \left( u'(c_t)e^{-\delta t} \right) /dt = -\eta(c) \frac{1}{c} \frac{dc}{dt} - \delta = -\eta g - \delta
\]

where \( g = \frac{1}{c} \frac{dc}{dt} \).

Now consider the problem

\[
Max_c \int_0^\infty u(c_t)e^{-\delta t}dt, \quad c_t + \frac{dk}{dt} = f(k)
\]
where we are maximizing the integral of the utility of consumption, discounted at rate $\delta \geq 0$, subject to the constraint that consumption plus investment $dk/dt$ adds up to output $f(k)$ where $k$ is the capital stock and $f$ a strictly concave production function. This is called the Ramsey problem or the optimal growth problem.

To solve this problem we use a Hamiltonian:

$$H = u(c_t) e^{-\delta t} + \lambda_t e^{-\delta t} [f(k_t) - c_t]$$

where $\lambda_t e^{-\delta t}$ is a time-varying shadow price, $c_t$ is called the control variable and $k_t$ the state variable, and the expression multiplied by the shadow price is the rate of change of the state variable.

First order conditions - necessary conditions - for a path of $c_t, k_t$ to solve this problem are that

$$\frac{\partial H}{\partial c_t} = 0 \forall t, \quad \frac{d}{dt} \left[ \lambda_t e^{-\delta t} \right] = -\frac{\partial H}{\partial k_t} \forall t$$

If all functions are concave these conditions are not only necessary but also sufficient, plus one additional technical condition known as a transversality condition.

Applying these conditions gives

$$u'(c_t) = \lambda_t$$

and

$$\frac{d\lambda_t}{dt} - \delta \lambda_t = -\lambda_t f'(k_t), \quad \frac{d\lambda_t}{dt} = 0 \iff \delta = f'(k)$$

from which

$$\eta g + \delta = f'(k_t)$$

So the return on capital - $f'(k)$ - has to equal the elasticity of MU times the growth rate of consumption plus the discount rate. And the LHS here is the rate of change of the marginal value of consumption.

### 12 State-Dependent Preferences

So far we have assumed that lotteries deliver money and that preferences are over amounts of money, with no other characteristics mattering. It may be
however that the circumstance under which money is delivered matter: for example, money if you have lost your house in an earthquake or hurricane may be more valuable than money if you just won the lottery. An umbrella is much more useful if it is about to rain than on a hot dry day. We will call circumstances such as whether it is dry or raining, or whether there is an earthquake or hurricane, “states of nature.” $S$ is the set of all possible states of nature (earthquakes, hurricanes, dry, wet, .... ), assumed to be finite, non-intersecting and exhaustive, and $s \in S$ is a particular state (earthquake etc). The probability of $s$ occurring is $\pi_s$.

**Definition 34.** A random variable is a function $g : S \rightarrow R_+$ that maps states into monetary outcomes.

Every random variable $g(.)$ gives rise to a money lottery described by the distribution function $F(.)$ where $F(x) = \sum_{s \in s | g(s) \leq x} \pi_s$ for all $x$. A random variable can now be represented by a vector of the monetary payoffs it gives in each of the states in $S$, denoted $(x_1, ..., x_S)$ where we are using $S$ also to denote the number of states in $S$. The set of all random variable is now $R^S_+$. In this framework the primitive concept is a preference ordering over the set of all random variables, $R^S_+$. Such a preference can be represented by an expected utility function as before, with one difference: the utility function can now depend on the state of nature. So we have

**Definition 35.** The preference relation $\succeq$ on $R^S_+$ has an expected utility representation if for every $s \in S$ there is a function $u_s : R_+ \rightarrow R$ such that for any $(x_1, ..., x_S)$ and $(x'_1, ..., x'_S) \in R^S_+$, $(x_1, ..., x_S) \succeq (x'_1, ..., x'_S)$ iff $\sum_s \pi_s u_s(x_s) \geq \sum_s \pi_s u_s(x'_s)$.

Consider a world with two states only, $s_1, s_2$. Let $x_1, x_2$ be monetary amounts delivered in each state. Then if $x_1 = x_2$ we have the same payoff in each state and there is no uncertainty about the payoff. The marginal rate of substitution between $x_1, x_2$ is now $\pi_1 u'_1(x) / \pi_2 u'_2(x)$ and if the function is state-independent then this reduces to the ratio of the probabilities. Generally it depends on probabilities and preferences.

With state-dependent preferences it is no longer the case that a risk-averse person will always purchase actuarially fair insurance. Insurance that pays $p_i$ in state $i$ is actuarially fair if $\pi_1 p_1 + \pi_2 p_2 = 0$, $p_2 = -p_1 \frac{\pi_1}{\pi_2}$. This means that the budget line for insurance has the slope $-\pi_1/\pi_2$, and this is the slope of an indifference curve on the risk free line only if preferences are not state-dependent.
13 Subjective Probabilities

Von Neumann and Morgenstern took probabilities as given, objective, and deduced the existence of preferences such that agents maximize the expectation of utility at these probabilities.

de Finetti, an Italian probabilist, worked the other way, and derived probabilities from agents' behavior. He said: suppose you have to set the price of a promise to pay $1 if there was life on Mars one billion years ago, and nothing otherwise. Your opponent can then choose which side of this bet you are on and which side he is on. de Finetti defined the price you set for this promise as your subjective probability of there having been life on Mars one billion years ago. Here we have deduced probabilities from observed behavior. They are not based on experimental evidence and can clearly differ from person to person.

Savage (Foundations of Statistics 1953) developed a theory of choice under uncertainty in which both preferences and probabilities are deduced from behavior. He postulated a larger and more demanding set of axioms, and in exchange for this greater complexity proved more - that people under these conditions will behave as if they have both preferences and personal probabilities and maximize their expected utilities using these preferences and probabilities.

Savage’s framework is as follows. Primitive concepts are states and outcomes. The set of states \( s \in S \) is an exhaustive list of all scenarios that might unfold. Knowing which state occurs resolves all uncertainty. An event is any subset \( A \subset S \). The set of outcomes is \( X \), typical member \( x \in X \). An outcome specifies everything that affects the chooser's well-being.

The objects of choice are acts (policies, strategies), which are functions from states to outcomes, and acts are denoted \( f \in F, f : S \rightarrow X \). The state is uncertain and so not known when the act is chosen, but the actor does know that if the state is \( s \) then the outcome is \( f(s) \).

Acts whose payoffs do not depend on the state of the world \( s \) are constant functions in \( F \). We will use the notation \( x \in F \) to indicate the constant function in \( F \) whose outcome is always equal to \( x \in X \). Suppose \( f, g \) are two acts and \( A \) is an event: then we define a new act by

\[
    f_A^g(s) = \begin{cases} 
    g(s), & s \in A, \\
    f(s), & s \in A^c
    \end{cases}
\]

Intuitively this is \( f \) but replaced on \( A \) by \( g \).
13.1 Savage’s Axioms

**Axiom P1.** Preferences are a complete transitive relation on $F$.

**Axiom P2.** Preferences between two acts $f, g$ depend only on the values of $f, g$ where they differ.

Let $A$ be an event and $A^c$ its complement. Suppose $f, g$ are equal if $A$ does not occur, that is on $A^c$, so they differ only on $A$. Alter $f, g$ only on $A^c$ to get $f', g'$ such that $f(s) = f'(s), g(s) = g'(s), s \in A$. So there is no change on $A$. They are still equal off $A$, though not necessarily to $f, g$: $f(s) = g(s), f'(s) = g'(s), s \in A^c$.

Then P2 requires:

$$f \succeq g \iff f' \succeq g'$$

This is sometimes written $f \succeq_A g$, $f$ preferred or indifferent to $g$ given $A$. This is often referred to as Savage’s “sure thing principle.”

**Axiom P3.** If you take an act that guarantees an outcome $x$ on an event $A$ and you change it on $A$ from $x$ to another outcome $y$, the preference between the two acts should follow the preference between the two outcomes. Formally let $f_A^x$ be an act that produces $x$ for every state in $A$. A null event is, roughly, one that is thought to be impossible. Formally an event $A$ is null if whenever two acts yield the same outcome off $A$ they are ranked as equivalent. For every act $f \in F$, every non-null event $A \subset S$, and $x, y \in X$,

$$x \succeq y \iff f_A^x \succeq f_A^y$$

Here $x, y$ can also be interpreted as the acts that yield $x, y$ respectively in every state. This is a monotonicity assumption. Another interpretation is that rankings should be independent of the events with which they are associated (note the requirement that this hold for every non-null event $A$).

**Axiom P4.** For every $A, B \subset S$ and every $x, y, z, w \in X$ with $x \succ y \& z \succ w$,

$$y_A^x \succeq y_B^x \iff w_A^z \succeq w_B^z$$

This is an axiom about probabilities: presumably $y_A^x \succeq y_B^x$ means that you think event $A$ is more likely than event $B$.

**Axiom P5.** There are $f, g$ such that $f \succ g$. 

66
**Axiom P6.** For every \( f, g, h \in F \) with \( f \succ g \) there exists a partition of \( S \) [a collection of pairwise disjoint events whose union is \( S \)] denoted \( \{A_1, A_2, \ldots, A_n\} \) such that for every \( i \)

\[
f_{A_i}^h \succ g & \& f \succ g_{A_i}^h
\]

This is roughly like a continuity assumption, but it is hard to state continuity in Savage’s framework.

**Axiom P7.** Consider acts \( f, g \in F \) and an event \( A \subset S \). If for every \( s \in S, f \succeq_A g(s) \) then \( f \succeq_A g \), and if for every \( s \in A, g(s) \succeq_A f \), then \( g \succeq_A f \).

**Proposition 32.** [Savage] Assume that \( X \) is finite. Then \( \succeq \) satisfies P1 to P6 if and only if there exists a probability measure \( \mu \) on states \( S \) and a non-constant utility function \( u : X \to \mathbb{R} \) such that for every \( f, g \in F \),

\[
f \succeq g \iff \int_S u(f(s)) \, d\mu(s) \geq \int_S u(g(s)) \, d\mu(s)
\]

Furthermore \( \mu \) is unique and \( u \) is unique up to positive linear transformations.

As a generalization we also have the same result for infinite state spaces if we use P7:

**Proposition 33.** [Savage] \( \succeq \) satisfies P1 to P7 if and only if there exists a probability measure \( \mu \) on states \( S \) and a non-constant utility function \( u : X \to \mathbb{R} \) such that for every \( f, g \in F \),

\[
f \succeq g \iff \int_S u(f(s)) \, d\mu(s) \geq \int_S u(g(s)) \, d\mu(s)
\]

Furthermore \( \mu \) is unique and \( u \) is unique up to positive linear transformations.

These results look at first sight like the von Neumann Morgenstern result, but in fact they are far stronger.

We use this theorem just as we use the vNM one, and can go through the same mechanisms with it, but the applicability is greater. These axioms produce a preference representation with a separation of preferences (utility function) from beliefs (probabilities). (Are preferences and beliefs really separate?)

An illustration of the sure thing principle, P2: here are four bets.
1. If horse A wins you get a trip to Paris, and otherwise you get trip to Rome

2. If horse A wins you get a trip to London and otherwise a trip to Rome

3. If horse A wins you get a trip to Paris and otherwise a trip to Los Angeles

4. If horse A wins you get a trip to London and otherwise a trip to Los Angeles

Clearly 1 and 2 are the same if A loses. Generally your choice will depend on preferences and beliefs or probabilities, but presumably the chance of A winning is the same in each case, so the choice depends on your preferences between Paris and London. The same is true for 3 and 4, and axiom P2 requires $1 \preceq 2 \iff 3 \preceq 4$. If two acts are equal on a given event, it does not matter what they are equal to. So it doesn’t matter if when the horse loses you get Rome or LA.

### 14 Non-Expected Utility Approaches

Suppose agents cannot derive subjective probabilities: they have no basis at all for assigning probabilities to events. An example is the probability that there is life in the universe on an planet other than ours within $10^{12}$ light years from us. How can we then describe rational choice under uncertainty? Or in the case of the Ellsberg paradox, where the numbers of black and yellow balls are unknown. Here is another version of the Ellsberg paradox.

There are two urns, each with 100 balls. Urn 1 has 50 black and 50 red. The numbers of red and black in urn 2 are not known. You are asked whether you would rather bet on a red ball being withdrawn from urn 1 than urn 2: most people reply yes. Then you are asked whether you would rather bet on a black ball being withdrawn from 1 than 2: again most people answer yes. But this is inconsistent with probabilistic reasoning. If you would sooner bet on a red ball being taken from urn 1 than urn 2 you must believe that it is more likely that the red ball will be taken from 1 than 2: the probability of its being taken from 1 is 0.5 so the probability of its coming from 2 is less than 0.5. Similarly for a black ball: if you prefer to bet on urn 1 then you must think a black ball is more likely from urn 1 than from urn 2, and it

68
chance from urn 1 is 0.5 and so it is less than this from urn 2. So the chances of both red and black from urn 2 are less than 0.5. But they have to sum to 1.

Here is a related example. You have to bet on the toss of a coin. There are two coins and you can choose which to toss. One has been tossed many times and came down heads 50% of them. The other has never been tested. Which would you rather bet on? In one case you know that the odds are 50/50: in the other case this is a reasonable assumption but you have no evidence. Most people would sooner bet on the tested coin.

In both of these examples people cannot quantify probabilities and stay away from bets involving unquantified risks.

14.1 MinMax Approaches

One set of approaches to these problems is to ignore probabilistic information. The classic approach of this type is the maxmin approach, due to Wald. Recall that $f$ is an act that maps states $S$ to outcomes $X$. We say that an act $f$ is preferred to an act $g$ if the worst outcome associated with the choice of $f$ is better than the worst outcome associated with $g$. Formally

$$f \succeq g \iff \min_{s \in S} f(s) \geq \min_{s \in S} g(s)$$

We then seek the act that is ranked best in this ordering:

$$\max_{f \in F} \min_{s \in S} f(s)$$

Another approach is the minmax regret approach, due to Savage. The regret associated with a state $s$ is the difference between the outcome according to the act chosen and the best possible act given that state:

$$r(s, g) = \max_{f \in F} f(s) - g(s)$$

The max regret for a given policy is the maximum of this for all possible states:

$$\max_{s \in S} r(s, g) = \max_{s \in S} \{ \max_{f \in F} f(s) - g(s) \}$$

and so is the worst shortfall you could have between actual and ideal outcomes under act $g$. The optimum policy then minimizes this over all acts:

$$\min_{g \in F} \max_{s \in S} r(s, g)$$
Both of these approaches neglect any probabilistic information available.

One approach that does take probabilistic information into account is to consider all probability distributions that are consistent with what we know, and use all of these in making a decision.

14.2 MaxMin Expected Utility

An approach due to Gilboa and Schmeidler (Maxmin Expected Utility with a Non-Unique Prior, Journal of Mathematical Economics, 1989, 18, 141-53) with the following axioms. The framework is as with Savage above.

Axiom 1. We have a complete transitive ordering over $F$.

Axiom 2. Continuity: For every $f, g, h \in F$ if $f \succ g \succ h$ then there exist $\alpha, \beta \in (0,1)$ such that

$$\alpha f + (1 - \alpha) h \succ g \succ \beta f + (1 - \beta) h$$

Axiom 3. Monotonicity: For every $f, g \in S$, $f(s) \succeq g(s) \forall s \in S \Rightarrow f \succeq g$

Axiom 4. Nontriviality: There exist $f, g \in X : f \succ g$

Axiom 5. Independence: For every $f, g \in F, \forall constant h \in F, \forall \alpha \in (0,1)$,

$$f \succeq g \Leftrightarrow \alpha f + (1 - \alpha) h \succeq \alpha g + (1 - \alpha) h$$

Axiom 6. Uncertainty Aversion. For every $f, g \in F, \forall \alpha \in (0,1), f \sim g \Rightarrow af + (1 - \alpha) g \succeq f$

Proposition 34. A preference satisfies the above axioms if and only if there exists a closed convex set of probabilities $C$ and a non-constant function $u : X \rightarrow R$ such that for every $f, g \in F$

$$f \succeq g \Leftrightarrow \min_{p \in C} \int_S u(f(s)) dp(s) \geq \min_{p \in C} \int_S u(g(s)) dp(s)$$

Furthermore in this case $C$ is unique and $u$ is unique up to a positive linear transformation.

What this theorem is saying, is: look at the probabilities that give the worst possible expected utility for each act, and evaluate the acts according to these probabilities. So evaluate an act by the probability that gives the minimal outcome and choose the best according to this ranking.
14.3 Smooth Ambiguity Aversion

In this case we again work with many probability distributions that are consistent with what we know. But rather than focussing only on the “worst” of them, in the sense of lowest expected utility, we give them all weights and take note of them all according to these weights. The weight attached to a distribution can be thought of as the subjective assessment of the chance of that probability distribution being the correct one. (Klibanoff, Marinacci and Mukerji, Decision-Making under Ambiguity, Econometrica, 2005, 73(6), 1848-1892)

We have many probability distributions $p$ over states. The first assumption is that for any probability over states $S$ there is a utility such that acts are ranked by the expectation of that utility:

**Axiom 1.** Let $p$ be a probability over states $S$. Then there exists $u : X \rightarrow R$ such that $f \succeq g \iff \int_S u(f(s)) \, dp \geq \int_S u(g(s)) \, dp$. This is the von Neumann Morgenstern theorem, taking the probability $p$ over states as given.

For each action $f \in F$ and each probability $p$ we now have an expected utility $E_p f = \int_S u(f(s)) \, dp$

**Axiom 2.** There exist weights $\pi(p) \geq 0$, $\int \pi(p) \, dp = 1$, and a function $\phi : R \rightarrow R$ such that

$$f \succeq g \iff \int_p \pi(p) \phi(E_p f) \, d\pi \geq \int_p \pi(p) \phi(E_p g) \, d\pi$$

In words this states that we prefer $f$ to $g$ if and only if the expectation of the function $\phi$ of the expected utilities according to the weights $\pi$ is greater for $f$ than for $g$. We can think of $\phi$ as a second order utility function - defined on expected utilities - and the weights as second order probabilities.

14.4 Examples

First look at the two-urn version of the Ellsberg paradox. Urn 1 has 100 balls, 50 red and 50 yellow. Urn 2 has also 100 red and yellow balls in unknown proportions. You are asked if you are interested in betting $10 on a red ball being drawn from urn 1 or urn 2. We consider the value of this bet firstly with linear utilities and then with concave utilities.

**Linear utilities:** the value of the bet on urn 1 is clearly $0.5 \times 10 + 0.5 \times 0 = 5$. 

71
With urn 2 all possible distributions of 100 balls between red and yellow are possible, and we can take either the mmu approach or the smooth ambiguity approach. Let \( p_n, n = 0, \ldots, 100 \) be the probability of choosing a red ball given that the number of red balls is \( n \). So \( p_n = n/100 \).

With mmu the worst possible distribution is that giving zero probability to choosing a red ball, \( p_0 \). In this case the value of the bet is zero, so the mmu approach values this bet at zero.

With smooth ambiguity we consider all probabilities \( p_0, \ldots, p_{100} \) and not just \( p_0 \). We give each probability \( p_n \) a weight or likelihood or second order probability \( \pi_n \). Then with linear functions \( u(x) = x, \phi(y) = y \) the value of the bet is

\[
\sum_{n=0}^{100} \pi_n p_n 10 = \sum_{n=0}^{100} \frac{n}{100} 10 \pi_n = \frac{1}{10} \sum_{n=0}^{100} n \pi_n
\]

and if we rank all numbers of red balls \( n \) as equally likely (\( \pi_n = \pi = \frac{1}{101} \)) then this is

\[
\frac{1}{10} \sum_{n=0}^{100} n \pi_n = \frac{1}{10} \sum_{n=0}^{100} \frac{n}{101} = 5
\]

which is the same as the value of the bet on urn 1.

Now assume ambiguity aversion but no risk aversion, so it is still the case that \( u(x) = x \). The value of a bet on urn 1 is just \( 0.5 \times 10 + 0.5 \times 0 = 5 \).

According to MMU the value of the bet on urn 2 is \( u(0) = 0 \).

Now consider the smooth ambiguity approach. If the number of red balls is \( n \), so the probability of red is \( n/100 \), the expected utility is \( \frac{n}{100} 10 + \frac{100-n}{100} 0 = \frac{n}{10} \). So overall the value of the bet is

\[
\sum_{n=0}^{100} \pi_n \phi\left( \frac{n}{10} \right)
\]

If \( \phi \) is linear this is the same as the value of a bet on urn 1. This is the case of no ambiguity aversion.

Consider instead the case of \( \phi(x) = x^{0.5} \), a strictly concave function. Then

\[
\sum_{n=0}^{100} \frac{1}{101} \phi\left( \frac{n}{10} \right) = \frac{1}{101} \sum_{n=0}^{100} \sqrt{\frac{n}{10}} = 2.1
\]

and the certainty equivalent of this bet is 4.4.
Here is another example. A model \( m_i \) is a mapping from acts \( f \in F \), which we take to be \( R^1 \) for this example, to probability distributions \( p(x) \) over outcomes \( x \in X \). (Think of this as a macroeconomic model or a model of the stock market, whose output is a distribution over possible outcomes, with the act being the choice of an interest rate or an investment level.) We want to choose the act that leads to the best outcomes but don’t know which model is the right one. So an act \( f \) gives a distribution over outcomes that depends on the model:

\[
m_i(f) = p(x \mid f, m_i)
\]

which is a distribution over \( x \) conditional on the act \( f \) and the model \( m_i \). The expected utility from an act contingent on model \( i \) being correct is therefore

\[
Eu(f \mid m_i) = \int u(x) \, dp(x \mid f, m_i)
\]

The mmu approach is to value each act \( f \) according to the model that gives the worst outcome:

\[
\min_{m_i} \int u(x) \, dp(x \mid f, m_i)
\]

and then maximize this value across acts:

\[
\max_f \min_{m_i} \int u(x) \, dp(x \mid f, m_i)
\]

The smooth ambiguity approach will assign second order likelihoods \( \pi_i \) or probabilities \( \pi_i \) to the models \( m_i \) and evaluate expected utilities by the concave function \( \phi \). So the problem is

\[
\max_f \sum_i \pi_i \phi(Eu(f \mid m_i))
\]

The FOCs for this are (assuming \( f \in R \))

\[
\sum_i \pi_i \phi'(Eu(f \mid m_i)) \, Eu'(f \mid m_i) = 0
\]

Define ambiguity-adjusted probabilities \( \pi_i' \) as

\[
\pi_i' = \frac{\pi_i \phi'(Eu(f \mid m_i))}{\sum_j \pi_j \phi'(Eu(f \mid m_j))}
\]
and divide the FOC through by the denominator of this expression to get a new way of stating the FOC:

$$\sum \pi'_i E u'(f \mid m) = 0$$

So the expected sum of the marginal expected payoffs from a change in the act $f$ must be zero, where the expectation is calculated at the ambiguity-adjusted probabilities. Because $\phi$ is concave and so $\phi'$ is decreasing, these adjusted probabilities give more weight to bad outcomes and less to good outcomes than the original second-order probabilities $\pi_i$.

**Problem**

A university has an endowment $W$ that it may invest in bonds $B$ or equity $E$. Each type of security may go up 10% or go down 10%. The distributions are not independent. The university has two financial advisers $X$ and $Y$ who give different estimates of the probabilities of the possible cases, and the university cannot tell which if either is correct. For adviser $X$ these probabilities are $x_{ij}$, and for $Y$ they are $y_{ij}$. The university evaluates outcomes according to a concave utility function $U(P)$ where $P$ is the financial payoff. Formulate the university’s investment problem according to the MaxMin Expected Utility approach and the Smooth Ambiguity approach.

**Answer**

Here is the table of possible outcomes and the probabilities that advisor $X$ assigns to them: for advisor $Y$ replace $x_{ij}$ by $y_{ij}$.

<table>
<thead>
<tr>
<th>B↓/E→</th>
<th>+10%</th>
<th>-10%</th>
</tr>
</thead>
<tbody>
<tr>
<td>+10%</td>
<td>$x_{11}$</td>
<td>$x_{12}$</td>
</tr>
<tr>
<td>-10%</td>
<td>$x_{21}$</td>
<td>$x_{22}$</td>
</tr>
</tbody>
</table>

The investment in equities is $eW$ and that in bonds is $(1 - e)W$. So the expected utility according to advisor $X$ is

$$EU_x = x_{11}U [1.1eW + 1.1 (1 - e) W] + x_{12}U [0.9eW + 1.1 (1 - e) W] + x_{21}U [1.1eW + 0.9 (1 - e) W] + x_{22}U [0.9eW + 0.9 (1 - e) W]$$
This can be simplified to

$$EU_x = K_x + x_{12}U\left[ W (1.1 - 0.2e) \right] + x_{21}U\left[ W (0.9 + 0.2e) \right]$$

where $K_x = x_{11}U[1.1W] + x_{22}U[0.9W]$

Note that

$e = 0 \Rightarrow EU_x = K_x + x_{12}U[1.1W] + x_{21}U[0.9W]$

$e = 1 \Rightarrow EU_x = K_x + x_{12}U[0.9W] + x_{21}U[1.1W]$

and

$$\frac{\partial EU_x}{\partial e} = 0.2 \left\{ x_{21}U'(C_{21}) - x_{12}U'(C_{12}) \right\}$$

where $C_{ij}$ is income in state $i, j$, and this derivative is positive when $e = 0$ and negative when $e = 1$.

For the **MaxMin Expected Utility** approach we need to find $EU_x(e)$ and $EU_y(e)$ for each value of $e$, pick the min,

$$V(e) = Min_{x,y} \{ EU_x(e), EU_y(e) \}$$

and then choose $e$ to maximize $V(e)$.

For the **Smooth Ambiguity Approach** we value a policy $e$ as follows: let $\pi_i$ be the probability the university assigns to advisor $i$ being right (a second order probability) and $\phi: R \to R$ be a concave increasing function. Then the objective is

$$V(e) = Max_{x,y} \{ \pi_x \phi(EU_x(e)) + \pi_y \phi(EU_y(e)) \}$$

To solve this maximization problem we choose $e^*$ so that

$$\frac{\partial V}{\partial e} = \pi_x \phi'(EU_x(e)) EU'_x(e) + \pi_y \phi'(EU_y(e)) EU'_y(e) = 0$$

Note that we can divide both sides of this equation by $\pi_x \phi'(EU_x(e)) + \pi_y \phi'(EU_y(e))$ giving

$$\frac{\pi_x \phi'(EU_x(e))}{\pi_x \phi'(EU_x(e)) + \pi_y \phi'(EU_y(e))} EU'_x(e) + \frac{\pi_y \phi'(EU_y(e))}{\pi_x \phi'(EU_x(e)) + \pi_y \phi'(EU_y(e))} EU'_y(e) = 0$$

which we can write as

$$\pi_x' EU'_x(e) + \pi_y' EU'_y(e) = 0$$
where \( \pi'_x, \pi'_y \) are ambiguity-adjusted second order probabilities.

\[
\pi'_x = \frac{\pi_x \phi'(EU_x(e))}{\pi_x \phi'(EU_x(e)) + \pi_y \phi'(EU_y(e))}, \quad \pi'_y = \frac{\pi_y \phi'(EU_y(e))}{\pi_x \phi'(EU_x(e)) + \pi_y \phi'(EU_y(e))}
\]

So the solution involves setting the expected marginal gain from shifting the portfolio equal to zero, where the expectation is taken via these ambiguity-adjusted second order probabilities. Note that if \( \phi \) is strictly concave then the ambiguity adjustment involves placing more weight on the bad outcome than with the initial second order probabilities. If \( \phi \) is linear there is no change.
15 Third Problem Set

Problem 1. Consider the following two lotteries: \( L = \$200 \) with probability 0.7, 0 with probability 0.3, and \( L' = \$1200 \) with probability 0.1 and \$0 with probability 0.9. Let \( x_L \) and \( x_{L'} \) be the sure amounts of money the individual finds indifferent to \( L \) and \( L' \) respectively. Show that if preferences are monotone, the individual must prefer \( L \) to \( L' \) if and only if \( x_L > x_{L'} \).

Problem 2. Consider the insurance problem studied in section 10.1, and show that if insurance is not actuarially fair \( (q > \pi) \) then the individual will not insure completely.

Problem 3. Show that if an individual has a Bernoulli utility function of the form
\[
u (x) = \beta x^2 + \gamma x
\]
then her utility from a distribution is determined by the mean and variance of the distribution and by these alone. Note: \( \beta < 0 \) for concavity of \( u \) and we limit the distribution to values no greater than \(-\gamma/2\beta\) as \( u \) is decreasing after this.

Problem 4. Assume that a firm is risk-neutral with respect to profits and that if there is uncertainty about prices then production choices are made after the resolution of this uncertainty. The firm faces a choice between two alternatives. In the first prices are uncertain. In the second they are certain and equal to the expected value of the uncertain case. Show that a firm that maximizes expected profits will prefer the first alternative to the second.

Problem 5. Suppose that an individual has a Bernoulli utility \( u (x) = x^{1/2} \).
1. Calculate the coefficients of absolute and relative risk aversion when \( x = 5 \)
2. Calculate the certainty equivalent for the gamble \((16, 4 : 0.5, 0.5)\)
3. Calculate the certainty equivalent for the gamble \((36, 16 : 0.5, 0.5)\)

Problem 6. Consider a lottery over monetary outcomes that pays \( x + \epsilon \) with probability 0.5 and \( x - \epsilon \) with probability 0.5. Compute the second derivative of this lottery’s certainty equivalent with respect to \( \epsilon \) and show that the limit of this derivative as \( \epsilon \to 0 \) is exactly \(-r_A (x)\).
Part IV
General Equilibrium

Next we study the interactions of firms and consumers through markets. Firms as before are characterized by production possibility sets: firm $i$ has production possibility set $Y_i \subset R^N$, and $y_i \in Y_i$ is production plan. There are $I$ firms. The sign convention as before is that inputs are negative and outputs positive. A price vector $p$ is an element of $R^N_+$ with $\sum_i p_i = 1, p_i \geq 0 \forall i$. Clearly profits are given by $\pi_i = p.y_i$. Firms seek to:

$$\max_{y_i \in Y_i} \{p.y_i\} = \pi_i$$

Consumers as before have preferences $\succeq_j$ on $X_j \subset R^N_+$ and a consumption vector is $x_j \in X_j$. There are $J$ consumers. Preferences are represented by an ordinal utility function $u_j : R^N \rightarrow R$. Consumers have endowments $w_j \in R^N$: $w_j$ is the vector of goods that individual $j$ owns and can either consume or sell. Typically it contains labor, which may be consumed as leisure or sold as work, and any other items that belong to the individual. Firms are owned by individuals: individual $j$ owns a fraction $\theta_{ji}$ of firm $i$, entitling her to this fraction of its profits. So the consumer choice problem is

$$\max_{x_j} u(x_j), \ p.x_j \leq p.w_j + \sum_i \theta_{ji}\pi_i$$

Here total spending power is the value of endowments plus the income from shareholdings. A set of consumption and production plans, one for each consumer and producer, is called an allocation.

**Definition 36.** Let $y_i^*, x_j^*$ be an allocation. We say this is feasible if

$$\sum_j x_j^* \leq \sum_j w_j + \sum_i y_i^*$$

that is if consumption is less than or equal to production plus endowments for each good or service.

Another important definition is that of Pareto efficiency (aka Pareto optimality)
Definition 37. An allocation \( y^*_i, x^*_j \) is **Pareto efficient** if it is feasible and there is no other feasible allocation \( \hat{y}_i, \hat{x}_j \) such that \( u_j(\hat{x}_j) \geq u_j(x^*_j) \) \( \forall j \), \( > u_j(x^*_j) \) some \( j \). In other words, there is no other feasible allocation where someone is better off and no-one is worse off.

We can also work with the Pareto ranking:

**Definition 38.** An allocation \( y^*_i, x^*_j \) is Pareto superior to another allocation \( \hat{y}_i, \hat{x}_j \) if \( u_j(x^*_j) \geq u_j(\hat{x}_j) \) \( \forall j \), \( \exists j : u_j(x^*_j) > u_j(\hat{x}_j) \). In words, someone is better off and no-one is worse off at the starred allocation. This is a partial ordering, like the vector ordering.

Next we introduce the idea of a competitive equilibrium, which is a set of prices, production and consumption plans such that firms are maximizing profits, consumers are maximizing utility, and all markets clear:

**Definition 39.** A **competitive equilibrium** is a price vector \( p^* \), a set of production plans \( y^*_i \) for each firm \( i \in I \), and a set of consumption vectors \( x^*_j \) one for each consumer \( j \in J \), such that consumers and producers maximize utilities and profits respectively and demand is less than or equal to supply:

1. \( \forall i, y^*_i \max p^* \cdot y, y \in Y_i \)
2. \( \forall j, x^*_j \max u(x), p^*x \leq p^*w + \sum_i \theta_{ji} \pi_i \)
3. \( \sum_j x^*_j \leq \sum_j w_j + \sum_i y^*_i \)

There are several questions one can ask about this concept. One is - when is it Pareto efficient? Another is - does such an equilibrium exist? We will investigate the first question extensively. Before tackling the general cases, we will look at a simple 2X2 case that can be studied geometrically, the case of two consumers trading two goods, with no production - a 2X2 exchange economy.

### 16 Edgeworth Box

Let consumers be \( a \) and \( b \), and goods 1 and 2. Consumption and endowment vectors are in \( R^2 \). The total endowment of good \( i \) is \( w_i = w_{ai} + w_{bi}, i = 1, 2 \). An **Edgeworth Box** (see figure 16.1) is a rectangle whose horizontal side
is of length $w_1$ and whose vertical side is $w_2$ long. The lower left corner is the origin for consumer $a$'s preferences and the top right corner is that for consumer $b$'s preferences, increasing to the south west. The budget line for a consumer is a line whose slope equals the price ratio, and which goes through that consumer’s endowment vector [because she can afford her endowment whatever the prices are]. Any point in this rectangle represents an allocation of the two goods between the two consumers: its coordinates relative to the normal lower left origin are the amounts allocated to consumer $a$, and the remaining amounts, which are the coordinates relative to the upper right origin, are the amounts allocated to consumer $b$. So in particular the initial endowments of the two consumers form a point in this box. A line through this point with slope equal to the price ratio gives the budget lines of both consumers.

By varying the slope of the budget line we can trace out the offer curves of each consumer, and check whether supply and demand match.
By looking at points where pairs indifference curves are tangential to each other, we can locate the Pareto efficient allocations, forming a set generally called the contract curve. We can also find the Pareto efficient points that are Pareto superior to the initial allocation - these are attractive points from a bargaining perspective.

Figure 16.2:

The equilibrium prices associated with any allocation are given by the slope of the budget line that goes through that allocation and is simultaneously tangent to indifference curves of \( a \) and \( b \); this is shown in figure 16.1. [Note - there may be more than one set of equilibrium prices associated with an initial allocation.] This is a point where demand and supply are equal and note that it is on the contract curve and so is Pareto efficient. So this geometric approach suggests that a competitive equilibrium is Pareto efficient. Figure 16.2 shows a configuration where demand and supply are not equated. The initial endowment is \( E \), and at the prices shown \( a \) wants to move to the point \( A \) where her indifference curve is tangent to the budget line, and likewise \( b \) wants to move to \( B \). So \( a \) wants to sell \( DE \) of good 1 and
buy CE of good 2, and b wants to buy GE of good 1 and sell FE of good 2. So we have demand exceeding supply for good 1 and vice versa for good 2.

Note that if the two agents’ preferences are identical and homothetic, then the contract curve will be a straight line along the diagonal of the box, and equilibrium prices will be independent of the initial allocation of endowments. Here’s the argument for the contract curve being the diagonal. On the contract curve both agents’ indifference curves have the same slope. This means they must both consume the two goods in the same proportions. This only happens on the diagonal.

17 The Theorems of Welfare Economics

We investigate the relationship between competitive equilibrium and Pareto efficiency.

Definition 40. An allocation \( x_j^*, y_i^* \) and a price vector \( p^* \) form a price equilibrium with transfers if there is an assignment of wealth levels \( W_j \) with \( \sum_j W_j = p^* \cdot \sum w_j + \sum_i p^* y_i^* \) [the value of the wealth is the sum of the value of endowments plus profits] such that

1. \( \forall i, y_i^* \max p^*.y, \ y \in Y_i \)
2. \( \forall j, x_j^* \max u_j (x_j), p^*.x_j \leq W_j \)
3. \( \sum_j x_j^* = \sum_j w_j + \sum_i y_i^* \)

In words, what we have here is a way of dividing endowments between people such that the allocation \( x_j^*, y_i^* \) forms a competitive equilibrium at prices \( p^* \).

Proposition 35. First theorem of welfare economics. If preferences are locally non-satiated, then if \( x_j^*, y_i^*, p^* \) is a price equilibrium with transfers, the allocation \( x_j^*, y_i^* \) is Pareto efficient. In particular any competitive equilibrium is Pareto efficient.

Proof. Preference maximization implies that anything that the consumer strictly prefers, is unaffordable to her. Formally,

\[ u_j (x_j) > u_j (x_j^*) \Rightarrow p^*.x_j > W_j \]
Local non-satiation implies an additional property:

$$u_j(x_j) \geq u_j(x_j^*) \Rightarrow p^* \cdot x_j \geq W_j$$

A consumption vector that is at least as good as $x_j^*$, costs at least as much. [This follows from local non-satiation: if there were a consumption vector that is at least as good and that costs less, then there would be another vector arbitrarily close to it that is better and also costs less, contradicting utility maximization.]

Let $x'_j, y'_i$ be an allocation that is Pareto superior to $x^*_j, y^*_i$, so that $u_j(x'_j) \geq u_j(x^*_j) \forall j$, $> u_j(x^*_j)$ some $j$. Then we must have

$$p^* \cdot x'_j \geq p^* \cdot x^*_j \forall j \& \exists j : p^* \cdot x'_j > p^* \cdot x^*_j$$

In this case we have that

$$\sum_j p^* \cdot x'_j > \sum_j W_j = \sum_j p^* \cdot \sum_j w_j + \sum_i p^* \cdot y^*_i$$

Because $y^*_i$ is profit-maximizing at prices $p^*$ for firm $i$, we know that

$$p^* \cdot \sum_j w_j + \sum_i p^* \cdot y^*_i \geq p^* \cdot \sum_j w_j + \sum_i p^* \cdot y'_i$$

and hence

$$\sum_j p^* \cdot x'_j > p^* \cdot \sum_j w_j + \sum_i p^* \cdot y'_i$$

so that the allocation $x'_j, y'_i$ cannot be feasible. Feasibility requires that

$$\sum_j x'_j = \sum_j w_j + \sum_i y'_i$$

which contradicts the previous inequality, taking the inner product with $p^*$ on both sides.

So the take-away here is that if consumers are maximizing utility and firms are maximizing profits, all facing the same prices, and markets clear [the allocation is feasible] then the allocation is Pareto efficient. All facing the same prices is crucial: it means that marginal rates of substitution in production and consumption are all the same.

83
Definition 41. An allocation $x^*_j, y^*_i$ and a price vector $p^*$ form a quasi-equilibrium with transfers if there is an assignment of wealth levels $W_j$ with $\sum_j W_j = p^*. \sum w_j + \sum_i p^*.y^*_i$ such that

1. $\forall i, y^*_i \max p^*.y, y \in Y_i$
2. $\forall j$, if $x_j \succ x^*_j \Rightarrow p^*.x_j \geq W_j$
3. $\sum_j x^*_j = \sum_j w_j + \sum_i y^*_i$

Note that the second condition here is different from that in definition 40: we are not asking for utility maximization, but that any preferred choice costs no less than the wealth level. Under some extra conditions this does imply utility maximization - as we will see. We will use the following result:

Proposition 36. Separating hyperplane theorem. Let $U, V$ be two non-empty closed bounded convex sets in $R^n$ with no interior points in common. Then $\exists p \neq 0 \in R^n$ and a scalar $z$ such that $u.p \geq z \forall u \in U, v.p \leq z \forall v \in V$.

Now we can prove

Proposition 37. Second theorem of welfare economics. Assume that all production sets $Y_i$ are convex and that all preferences are also convex and locally non-satiated. Then for every Pareto efficient allocation $(x^*_j, y^*_i)$ there is a price vector $p \neq 0$ such that $(p, x^*_j, y^*_i)$ form a quasi-equilibrium with transfers.

Proof. 1. Let $V_j = \{ x \in R^N : u_j(x) > u_j(x^*_j) \}$ and define $V = \sum_j V_j$. This is the set of aggregate allocations that can make everyone better off than at $x^*_j$. Also let $Y = \sum_i Y_i$. Clearly $V, V_j, Y$ are all convex sets.

2. Call $\sum_j w_j = w$, the aggregate endowment. The set $Y + w$, the aggregate production set translated by the aggregate endowment, is the set of all aggregate bundles available for consumption given the technology and endowments.

3. Note that $V \cap (Y + w) = \emptyset$. This is an implication of the Pareto efficiency of the allocation - if this intersection were non-empty then there would be a vector that is feasible (in $Y + w$) and can be used to give every consumer a greater utility level than $x^*_j$ (is in $V = \sum_j V_j$).

4. Next note that there is a vector $p \neq 0 \in R^N$ and a number $r$ such that $p.z > r \forall z \in V$ and $p.z \leq r \forall z \in Y + w$. This is just an application of
the separating hyperplane theorem: \( V \) and \( Y + w \) are convex sets with no interior points in common (we proved this above in 3).

5. Next note that \( u_j(x_j) \geq u_j(x_j^\ast) \Rightarrow p(\sum_j x_j) \geq r \). Why? By local non-satiation there is a vector \( \hat{x}_j \) arbitrarily close to \( x_j \) such that \( u_j(\hat{x}_j) > u_j(x_j) \) and so \( \hat{x}_j \in V_j \Rightarrow \sum_j \hat{x}_j \in V \), so \( \left( \sum_j \hat{x}_j \right) \geq r \), and taking the limit as \( \hat{x}_j \to x_j \) we have that \( \sum_j p.x_j \geq r \).

6. \( \sum_j x_j^\ast = \sum_j (w + \sum_i y_i^\ast) = r \). By 5 \( \sum_j x_j^\ast \geq r \), but in addition we know that \( \sum_j x_j^\ast = \sum_i y_i^\ast + w \in Y + w \) and so \( \sum_j x_j^\ast \leq r \), implying that in fact \( \sum_j x_j^\ast = r \). Since \( \sum_j x_j^\ast = w + \sum_i y_i^\ast \) we also know that \( \sum_j (w + \sum_i y_i^\ast) = r \).

7. For each firm consider an arbitrary \( y_i \in Y_i \). Clearly \( y_i + \sum_{h \neq i} y_h^* \in Y \), so that by separation (point 4 above)

\[
p \left( w + y_i + \sum_{h \neq i} y_h^* \right) \leq r = p \left( w + y_i^* + \sum_{h \neq i} y_h^* \right)
\]

Hence \( p.y_i \leq p.y_i^* \).

8. For any consumer \( u_j (x_j) > u_i (x_j^\ast) \Rightarrow p.x_j \geq p.x_j^\ast \). Assume \( u_j (x_j) > u_i (x_j^\ast) \). By 5 and 6 above we have

\[
p \left( x_j + \sum_{k \neq j} x_k^\ast \right) \geq r = p \left( x_j^\ast + \sum_{k \neq j} x_k^\ast \right)
\]

so \( p.x_j \geq p.x_j^\ast \).

9. The wealth levels \( p.x_j^\ast = W_j \) support \( (p, x_j^\ast, y_i^\ast) \) as a quasi-equilibrium with transfers. Conditions 1 and 2 of definition 41 follow from 7 and 8 above. Condition 3 follows from the feasibility of a Pareto efficient allocation. □

What is the distinction between a quasi-equilibrium and an equilibrium, and when are they the same? In a quasi-equilibrium, anything that is better, costs no less: in an equilibrium, it costs more. When might something that is better not cost more? When the price of a good in which you are not satiated is zero. In this case in \( R^2 \) the budget line is horizontal or vertical and a consumer may wish to go infinitely far along it but be limited by the amount of the good available - the total endowment.
Something similar could happen if the consumer has endowments only of goods that have zero prices, so effectively has zero income.

If every consumer has positive wealth at a quasi-equilibrium with transfers then it is a price equilibrium with transfers and so any Pareto efficient allocation can be supported as a price equilibrium with transfers.

**Proposition 38.** Suppose that at a quasi-equilibrium every consumer has strictly positive wealth $W_j > 0$. Then it follows that $x_j \succ_j x_j^*$ \Rightarrow p.x_j > W_j$ so that the quasi-equilibrium is a price equilibrium with transfers.

*Proof.* Suppose in contradiction that there is $x_j \succ_j x_j^*$ & $p.x_j = W_j$. There exists an $x_j'$ such that $p.x_j' < W_j$. But for all $a \in [0, 1)$, $p[x_j + (1 - a)x_j'] < W_j$. But if $a$ is close enough to 1, the continuity of $\succeq_j$ implies that $ax_j + (1 - a)x_j' \succ_j x_j^*$, contradicting point 2 of definition 41.
Problem 1.

Consider an Edgeworth box economy in which consumers have the Cobb-Douglas utility functions \( u_1(x_{11}, x_{21}) = x_{11}^{\alpha_1} x_{21}^{1-\alpha_1} \) and \( u_2(x_{21}, x_{22}) = x_{21}^{\beta_2} x_{22}^{1-\beta_2} \). Consumer \( i \)'s endowments are \((\omega_{1i}, \omega_{2i}) > 0\). Solve for the equilibrium price ratio and allocation. How do these change with a small change in \( \omega_{11} \)?

Problem 2

Give the mathematical formula for two preferences which lead to an Edgeworth box in which prices are independent of the initial allocation of endowments amongst the agents. Prove that in this case the prices are independent of the initial allocation. Assume that the total endowments of the two goods are equal.

Problem 3

Construct, and give the mathematical formulae for, an Edgeworth box in which there is more than one competitive equilibrium from some initial allocations. Hint: you might work with linear preferences.
19 Existence of General Equilibrium

We now know something about the welfare properties of a general equilibrium. But we don’t actually know if such an equilibrium exists. This is not a trivial question - it is easy to construct examples of economies where there is clearly no price vector at which all markets clear simultaneously. So this question needs some work. The main concept in working on this is the excess demand function. Recall definition 39: a competitive equilibrium is a set of consumption plans $x^*_j$, production plans $y^*_i$ and prices $p^*$ such that

1. $\forall i, y^*_i \max p^*.y, \ y \in Y_i$
2. $\forall j, \ x^*_j \max u_j(x_j), \ p^*.x_j \leq W_j$
3. $\sum_j x^*_j = \sum_j w_j + \sum_i y^*_i$

So firms are maximizing profits, consumers utility, and demand and supply balance. Now define for any price vector $z(p)$, the excess demand associated with that price:

$$z(p) = \sum x_j(p) - \sum w_j - \sum y_i(p)$$

where $y_i(p) = \text{ArgMax } p.y, y \in Y_i, x_j = \text{argmax } U_j(x), p.x \leq p.w_j + \sum \theta_{ji} \pi_i$. So $z(p)$ is the difference between demand and supply at prices $p$, and for an equilibrium we need this to be non-positive for all goods. It can be negative - supply greater than demand - for goods whose price is zero. So the question now is: does there exist a price $p^*$ such that $z(p^*) \leq 0$?

Note that $z(p)$ is a map from prices to commodity space. Prices can be considered as points in the simplex in $R^N, S^N = \{p \in R^N : p_l \geq 0, \sum p_l = 1\}$. We know that demand and supply functions are homogeneous of degree zero in prices so we can always scale prices to be in the simplex without changing excess demand.

Next we modify $z(p)$ to $z'(p)$ which is a map from the simplex to itself. Define the following function $z^+(p)$ on $S^N: z^+_l(p) = \text{Max } \{z_l(p), 0\}$. Note that $z^+$ is continuous and that

$$z^+(p).z(p) = \sum_l \text{Max } \{z_l, 0\} z_l = 0 \Rightarrow z(p) \leq 0$$

Now construct

$$a(p) = \sum_l [p_l + z^+_l(p)]$$
Clearly we have \( a(p) \geq 1 \forall p \) as \( \sum_i p_i = 1 \). Now define the continuous function from \( S^N \) to itself by

\[ f(p) = \frac{1}{a(p)} [p + z^+(p)] \]

This is a continuous function from a compact convex set \( S^N \) to itself and so has a fixed point \( p^* \) (Brower’s fixed point theorem) such that \( f(p^*) = p^* \). By Walras’ Law

\[ 0 = p^* z(p^*) = f(p^*) \cdot z(p^*) = \frac{1}{a(p^*)} [p^* + z^+(p^*)] \cdot z(p^*) = \]

\[ \frac{1}{a(p^*)} [p^* \cdot z(p^*) + z^+(p^*) \cdot z(p^*)] = \frac{1}{a(p^*)} z^+(p^*) \cdot z(p^*) \]

Therefore \( z^+(p^*) \cdot z(p^*) = 0 \), which from above means that \( z(p^*) \leq 0 \) as required. So we have proved that there is a price at which all markets clear, provided that all demand and supply functions are continuous functions, which requires that all production sets be strictly convex and all utilities be strictly quasi concave. We have:

**Proposition 39.** If all production sets in the economy are strictly convex and all utilities strictly quasi-concave then there exists a price vector \( p^* \) at which \( z(p^*) \leq 0 \), that is, at which all markets clear.

Now we move on to consider some situations where competitive equilibria are not efficient.

## 20 Public Goods

A **public good** is one that, if provided for one person, is provided for all, or for local public goods, for all in a group or location. The traditional textbook examples are law and order, public health, and defense. A more contemporary examples is air quality, which if improved for one person in a region is necessarily improved for all. Public goods are said to be non-excludable (the provider cannot exclude from consuming the good someone who does not pay for it) and non-rivalrous (one person’s consuming the good does not prevent another from doing so).

The traditional market mechanism does not work well for public goods, as the provider cannot ensure that everyone who benefits from these goods,
pays for them. For example, a group in New York city may decide to incur costs to make the air in NYC cleaner and less dangerous, and ask people to pay for this. But they have no way of ensuring that every who benefits, pays. People can “free ride,” enjoy the benefits without paying. This is why we refer to “public goods,” reflecting the fact that these are normally provided by the government, which has the ability to force people to pay via taxes.

<table>
<thead>
<tr>
<th>EXCLUDABLE/RIVAL</th>
<th>yes</th>
<th>no</th>
</tr>
</thead>
<tbody>
<tr>
<td>yes</td>
<td>private</td>
<td>Impure public (uncongested toll road)</td>
</tr>
<tr>
<td>no</td>
<td>Impure public (crowded road, fishery)</td>
<td>public (air quality)</td>
</tr>
</tbody>
</table>

20.1 A Simple Case

We begin by considering a public good that is either provided or not: it is discrete and its scale cannot be altered. Think of a bridge or a tunnel: it is either built or not built. There are two people denoted by \( i, i \in \{1, 2\} \). There is one private good as well as the public good, and \( i \)'s consumption of the private good is \( x_i \). The amount of the public good is \( G \), the cost of providing it is \( c \), and \( i \)'s wealth is \( w_i \). Each person has to choose how much to spend on the private good and the public good, and the budget constraint is \( x_i + g_i = w_i \) where \( g_i \) is the amount that \( i \) spends on the public good. Utilities are given by \( U_i(G, x_i) \).

The amount of the public good is either zero or one: \( G \in \{0, 1\} \). The rule for its provision is

\[
G = \begin{cases} 
1 & \text{if } g_1 + g_2 > c \\
0 & \text{if } g_1 + g_2 < c 
\end{cases}
\]

So utilities are

\[
\begin{cases} 
(a) & U_i(1, w_i - g_i) \text{ if } g_1 + g_2 > c \\
(b) & U_i(0, w_i) \text{ if } g_1 + g_2 < c
\end{cases}
\]

A natural question to ask is: when is it Pareto efficient to provide the public good? We are asking here when outcome \( (a) \) is Pareto superior to outcome \( (b) \) above. This requires that there exist \( g_1, g_2 \) with \( g_1 + g_2 > c \) and \( U_i(1, w_i - g_i) > U_i(0, w_i) \) \( i = 1, 2 \).

Define \( r_i \) as \( i \)'s willingness-to-pay (wtp) for the public good as follows:

\[
U_i(1, w_i - r_i) = U_i(0, w_i) \quad (20.1)
\]
So this is the maximum amount $i$ could pay for the public good and be no worse off because of its provision. Combining these statements we find that

$$
U_1(1, w_1 - g_1) > U_1(0, w_1) = U_1(1, w_1 - r_1)
$$

$$
U_2(1, w_2 - g_2) > U_2(0, w_2) = U_2(1, w_2 - r_2)
$$

and with a little reorganization this implies that

$$
r_1 + r_2 > c
$$

(20.2)

So it is efficient to provide the public good if the sum of everyone’s wtp exceeds the cost. And next we show the converse: if the sum of wtps exceeds the cost then it is efficient to provide the public good. To show this suppose that $r_1 + r_2 > c$ and then choose $g_i < r_i$ and $g_1 + g_2 \geq c$ (which we can do because the first inequality is strict) and $U_i(1, w_i - g_i) > U_i(0, w_i)$ (which is possible because $U_i(1, w_i - r_i) = U_i(0, w_i)$ and $g_i < r_i$). Then we have found that when $r_1 + r_2 > c$ it is feasible and Pareto efficient to provide the public good. Summarizing:

**Proposition 40.** It is Pareto efficient to provide the public good if and only if the sum of agent’s willingness-to-pay exceeds the cost of the public good.

### 20.1.1 Digression on Game Theory

The prisoners’ dilemma game is famous. Two members of a criminal gang A and B are arrested and held in solitary confinement. The prosecutor lacks evidence to convict them of a major crime though she believes them to be guilty of one. The only way she can get evidence is if one of them confesses. However she does have evidence to convict them of a minor crime. Each gang member has two choices: he can confess to the major crime or he can remain silent. The outcomes are: if both remain silent then both get one year in prison: if both confess then both get two years, if one confesses and the other remains silent then the silent one gets three years and the other goes free. So we have the following table:

<table>
<thead>
<tr>
<th>A\B</th>
<th>Silent</th>
<th>Confess</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silent</td>
<td>1/1</td>
<td>3/0</td>
</tr>
<tr>
<td>Confess</td>
<td>0/3</td>
<td>2/2</td>
</tr>
</tbody>
</table>

where $x/y$ indicates that A gets $x$ years and B $y$ years. Confessing the the best strategy for either. Consider A’s choices: if B is silent then if A
confesses he goes free and if he remains silent he gets 1 year. So confessing beats remaining silent. And if B confesses, then if A confesses he gets 2 years and if he remains silent he gets 3, so again confessing beats remaining silent. The same applies to B’s choices. So both will confess, even though both would be better off if both remained silent. Many people see this as a parable about the irrationality of individual choices in a collective context. Confessing is a dominant strategy in this game - the best strategy whatever the other player does. Not many games have dominant strategies.

Next we look at the Nash equilibrium (watch Russel Crow in A Beautiful Mind, or read the book). We have players indexed by \( i, i = 1, 2, \ldots, I \) in a game. Each can choose a move (strategy) \( s_i \) from a set of possible strategies \( S_i \), and the utility of each player depends on the strategies chosen by all: \( U_i = U_i(s_1, s_2, \ldots, s_I) \). Let \( s_{-i} = (s_1, s_2, \ldots, s_{i-1}, s_{i+1}, s_{i+2}, \ldots, s_I) \). Then \( (s^*_i)_{i=1,\ldots,I} \) form a Nash Equilibrium if \( \forall i, s_i^* \max_{s_i} U_i(s_i, s_{-i}) \). So each person’s strategy is her best response to the strategies chosen by all others: each is maximizing her payoff given what the others are doing. If utilities are continuous and have convex upper contour sets and the strategy sets \( S_i \) and convex closed and bounded then we can show that a Nash equilibrium exists. Agent \( i \)'s reaction function \( f_i(s_{-i}) \) is a map from \( s_{-i} \) to \( s_i \) defined as follows:

\[
f_i(s_{-i}) = \max_{s_i} U_i(s_i, s_{-i})
\]

For each set of choices by all other agents, it shows agent \( i \)'s best response. We can see that a Nash equilibrium is a set of strategies \( s_i^* \) such that \( \forall i, s_i^* = f_i(s_{-i}^*) \). So it is the intersection of all reaction functions. In the case of two agents we have \( s_1^* = f_1(s_2^*) \) & \( s_2^* = f_2(s_1^*) \).

### 20.2 Private Provision in the Simple Case

We can use these ideas from game theory to analyze the private provision of public goods. Let the wtp's \( r_i = 100 \) and the cost \( c = 150 \). So it is Pareto efficient to provide the public good. We assume that each agent decides independently of the other whether or not to pay for the public good (i.e. to contribute 150 to its production). There are four cases: both buy, both don’t buy, 1 buys and 2 doesn’t and vice versa. The following table gives the payoffs in these four cases:
1↓/2→ | Buy | Not buy
---|---|---
Buy | -50/-50 | -50/100
Not buy | 100/-50 | 0/0

If both buy then they both lose 50: if neither buys they lose nothing, and if 2 buys and 1 doesn’t then 2 loses 50 and 1 gains 100, her wtp for the public good. Clearly not buying is a dominant strategy. So if both are asked to pay the full cost of the good then neither will do so.

Next consider selling the public good. Assume it is produced by a firm that seeks to maximize profits and that it can sell to each person a person-specific price $p_i$. Set $p_i = r_i$, so the good is sold to each person at their wtp. Then clearly the provision is Pareto efficient (the firm provides the good iff the sum of wtps exceeds the cost) and the firm maximizes its profits.

Finally consider a case in which each person offers to meet some of the cost of the public good: each offers $o_i$, and the good will be provided iff $o_1 + o_2 > c$. Assume that the provision of the good is Pareto efficient. Clearly each person will set $o_i \leq r_i$, that is no one will offer more than their wtp. Assume that $r_1 + r_2 > c$ and that the first person offers $o_1 < c$. Let $r_2 > c - o_1$, so person two values the good at more than the shortfall. How much should she offer? If she offers $o_2 < c - o_1$ then the good will not be provided and her utility is $U_2(0, w_2)$: if she offers $o_2^*(o_1) = c - o_1$ then her utility is $U_2(G, w_2 - (c - o_1)) > U_2(G, r_2) > U_2(0, w_2)$ by Pareto efficiency. So this maximizes her utility and is her best response to $o_1$. By exactly the same arguments if person 2 offers $o_2^*(o_1)$ then person one’s best response is just $o_1$ and the pair $(o_1, o_2^*(o_1))$ form a Nash equilibrium. We can investigate what these look like geometrically. Figure 20.1 shows this for the case when $c = 150$, $r_1 > 150 \forall i$. 
Figure 20.1:

The cost of the public good is 150 and both agents are willing to pay more than this. So if agent 1 offers $x$, then agent 2’s best response is to offer $150 - x$. In this figure each agent’s reaction function contains the line from $(150, 0)$ to $(0, 150)$ and so this is the set of Nash equilibria.
In the case shown in figure 20.2, reaction functions are more complex. Each values the good at less than its cost but together they value it at more than cost so it is efficient for it to be provided. If agent 1 offers $x$, then $2’s$ best response is $150 - x$ provided that $150 - x < 100$, or $x > 50$. If $x < 50$ then the best response is zero and likewise if $s \geq 105$. Agent $2’s$ is the same with the axes interchanged, and the set of Nash equilibria is $\{o_1, o_2 : o_1 + o_2 = 150, o_i \geq 50, i = 1, 2 \text{ and } [0,0]\}$.

### 20.3 The General Case

First a result on characterizing Pareto efficient allocations. The basic proposition is that we can characterize PE allocations as allocations that maximize a weighted sum of utilities.

**Proposition 41.** Suppose that the consumption vectors $x^*_j$ maximize the weighted utility sum $\sum_j \alpha_j U_j(x_j), \sum_j x_j \in \left( \sum_i Y_i + \sum_j w_j \right)$ where $\alpha_j \geq 0 \forall j$. Then $x^*_j$ are PE.
Proof. Suppose the $x^*_j$ are not PE. Then there exists an alternative set of consumption vectors $x'_j$ which are feasible ($\sum_j x'_j \in \left( \sum_i Y_i + \sum_j w_j \right)$) and such that $U_j (x'_j) \geq U_j (x^*_j) \forall j \& \exists j : U_j (x'_j) > U_j (x^*_j)$. In this case $\sum_j U_j (x'_j) > \sum_j U_j (x^*_j)$ a contradiction.

Let $c_j \in R$ be consumption of a normal, private, good, and $g \in R$ be the level of provision of a public good. For each individual $j$ utility depends on the consumption of both: $U_j (c_j, g)$ where $g$ being the consumption of the public good is the same for all, but people can of course choose different levels of the private goods. We assume $U_j$ to be concave. We suppose that each person has a budget that can be divided between the regular consumption goods and contributing to the provision of the public good. The total amount of the public good provided is a function of the total amount contributed by all individuals: $g = f(\sum_k g_k)$ where $g_j$ is $j$'s contribution to the public good and $f$ is concave. Hence the individual optimization problem is (setting the price of the private good equal to one)

$$\text{Max}_{c_j, g_j} U_j (c_j, g), \quad g = f(\sum_k g_k), \quad c_j = W_j - g_j$$

In this problem the agent is optimizing over her choice of consumption of the private good and contribution to the public good, taking as given the contributions that she thinks others are making ($g_k, k \neq j$) as in a Nash non-cooperative equilibrium, giving rise to the Lagrangian

$$L = U_j \left( c_j, f(\sum_k g_k) \right) + \lambda_j \left( W_j - g_j - c_j \right)$$

and the first order conditions are

$$\frac{\partial U_j}{\partial c_j} = \lambda_j, \quad \frac{\partial U_j}{\partial g} f' = \lambda_j \forall j$$

so that

$$\frac{\partial U_j}{\partial c_j} / \lambda_j = f' \quad \text{or} \quad \frac{\partial U_j}{\partial g} / f' = \frac{1}{\partial c_j}$$

This last expression says that the marginal rate of substitution between the public and private good should equal the marginal rate of transformation between them.
Now look at the socially efficient allocation between public and private goods. Consider the problem

$$\text{Max } \sum_j U_j(c_j, g), \ g = f \left( \sum_k g_k \right), \ \sum_j c_j = \sum W_j - \sum_k g_k$$

which leads to a Pareto efficient allocation with the public good. The Lagrangean is

$$L = \sum_j U_j \left( c_j, f \left( \sum_k g_k \right) \right) + \lambda \left\{ \sum_j W_j - \sum_k g_k - \sum_j g_j \right\}$$

and first order conditions are

$$\frac{\partial U_j}{\partial c_j} = \lambda, \ f' \sum_j \frac{\partial U_j}{\partial g} = \lambda$$

implying that

$$\sum_j \frac{\partial U_j}{\partial g} = \frac{\partial U_j}{\partial c_j} \frac{1}{f'} \text{ or that } \frac{\partial U_j}{\partial g} = \frac{\partial U_j}{\partial c_j} \frac{1}{f'} - \sum_{l \neq j} \frac{\partial U_l}{\partial g}$$

We can also write

$$\sum_j \frac{\partial U_j}{\partial g} \frac{\partial U_j}{\partial c_j} = \frac{1}{f'}$$

This is known as the Bowen-Lindahl-Samuelson formula for the optimal provision of a public good. Note the difference between the first order condition that the individual chooses

$$\frac{\partial U_j}{\partial g} = \frac{\partial U_j}{\partial c_j} \frac{1}{f'}$$

and that which is Pareto efficient

$$\sum_j \frac{\partial U_j}{\partial g} \frac{\partial U_j}{\partial c_j} = \frac{1}{f'}$$. Individual choices do not lead to an efficient outcome in this case. In fact it is easy to see that individual choices under-provide the public good relative to a Pareto efficient outcome, for clearly

$$\sum_j \frac{\partial U_j}{\partial g} \frac{\partial U_j}{\partial c_j} > \frac{\partial U_j}{\partial g} \frac{\partial U_j}{\partial c_j}$$

and if $f$ is strictly concave then this implies that the optimal level of provision is greater than the private level.

Now let’s look into having a market for the public good in which different people pay different prices and the good is provided by a profit-maximizing firm. So person $j$ pays price $p_j$ for the public good (the price of the private
good is one), and everyone pays the provider of the public good, the production of which uses as an input the private good. So the individual problem is

\[
Max_{c_j,g} U_j(c_j,g), \ W_j - c_j - p_j g = 0
\]

and the first order conditions are

\[
\frac{\partial U_j}{\partial c_j} = \lambda_j, \ \frac{\partial U_j}{\partial g} = \lambda_j p_j, \ \text{so} \ \frac{\partial U_j}{\partial c_j} / \frac{\partial U_j}{\partial g} = p_j
\]

and the problem for the firm producing the public good is

\[
Max g \sum_k p_k - z \text{ where } g = f(z)
\]

and the FOCs are

\[
f' \sum_k p_k = 1 \text{ or } \sum_k p_k = \frac{1}{f'}
\]

Combining both sets of FOCs we see that

\[
\sum_j \frac{\partial U_j}{\partial c_j} / \frac{\partial U_j}{\partial g} = \frac{1}{f'}
\]

which is the condition needed for Pareto efficiency.

**Example**

Let \( u(c_j,g) = \gamma \log(g) + \log(c_j) \). Let \( J \) be the total number of people, \( W \) the total endowment and \( W/J \) the share of each person. Assume the public good is produced from the private according to \( g(z) = z \) where \( z \) is the amount of the private good allocated in total to the production of the public good.

A Pareto efficient allocation is the solution to

\[
Max \sum_j \{ \gamma \log(g) + \log(c_j) \}, \ g + \sum_j c_j = W
\]

The Lagrangean is

\[
L = \sum_j \{ \gamma \log(g) + \log(c_j) \} + \lambda \left\{ W - g - \sum_j c_j \right\}
\]
and the FOCs are
\[ J \frac{\gamma}{g} = \lambda, \quad \frac{1}{c_j} = \lambda \]
from which it follows that the efficient allocation satisfies
\[ c_j = \frac{W}{J(1 + \gamma)}, \quad g = \frac{W\gamma}{1 + \gamma} \]

The typical individual solves the problem
\[ \text{Max}\ \{\gamma \log(g) + \log(c_j)\}, \quad g = z_j + \sum_{k \neq j} z_k, \quad c_j = \frac{W}{J} - z_j \]
Substituting into the utility function we find the FOCs are
\[ \frac{\gamma}{z_j + \sum_{k \neq j} z_k} = \frac{1}{\frac{W}{J} - z_j} \]
or
\[ \frac{W\gamma}{J} = z_j (1 + \gamma) + \sum_{k \neq j} z_k \]
and summing over \( j \) gives
\[ W\gamma = \sum_j z_j (1 + \gamma) + \sum_j \sum_{k \neq j} z_k \quad \text{or} \quad W\gamma = (J + \gamma) \sum_j z_j \]
so that
\[ g = \frac{\gamma W}{J + \gamma}, \quad c_j = \frac{W}{J + \gamma} \]
which is the same as the efficient allocation only if \( J = 1 \).

## 21 External Effects
These occur whenever the consumption/production of one person/firm affects the welfare/profits of another person/firm. Let \( c_j \) be the consumption vector of agent \( j \), and \( c_{-j} \) be the vector of consumption vectors of all agents other than \( j \). The for each agent utility takes the form \( U_j(c_j, c_{-j}) \). Individual endowments are \( W_j \). A social optimum is the solution to
\[ \text{Max} \ \sum_j a_j U_j(c_j, c_{-j}), \quad \sum_j p c_j = \sum_j W_j \]
where the $a_j \geq 0$ are welfare weights. The Lagrangean is

$$L = \sum_j a_j U_j (c_j, c_{-j}) + \lambda \left( \sum_j p_c c_j - \sum_j W_j \right)$$

and the FOCs are

$$\frac{1}{\lambda} \frac{\partial U_j}{\partial c_{j,l}} = p_l - \frac{1}{\lambda} \sum_{k \neq j} a_k \frac{\partial U_k}{\partial c_{j,l}}$$

Noting that if the external effects are harmful the terms $\frac{\partial U_k}{\partial c_{j,l}}$ are negative, and we can think of the second term on the RHS as a tax, which corrects for the external costs by adding them to the market prices of goods.

The private optimum is the solution to

$$\text{Max } U_j (c_j, c_{-j}) \text{ p}_c c_j = W_j$$

and clearly the FOCs are

$$\frac{1}{\lambda} \frac{\partial U_j}{\partial c_{j,l}} = p_l$$

which differ from the social optimum by the term $\frac{1}{\lambda} \sum_{k \neq j} a_k \frac{\partial U_k}{\partial c_{j,l}}$. Hence if this amount is added to the price $p_l$ for individual $j$, the private and social FOCs will coincide. Note that the tax is in principle person-specific. These taxes are known as “Pigovian taxes” after Arthur Pigou. Pollution taxes (carbon taxes) are examples of Pigovian taxes, and cap and trade systems are also systems for adding external costs to the prices that agents face.

22 Common Property Resources

Common property resources are resources to which all have equal access. A classic example is fisheries, though ground water is also a common property resource. Assume that the total production from a common property resource $Y$ is a function $F$ of the total inputs applied to the resource $X$, $Y = F(X)$, where $X = \sum_i x_i$, and $x_i$ is the input applied by person $i$. So $Y$ could be the catch from a fishery and $x_i$ the number of vessel-hours applied to the fishery by agent $i$. Alternatively $Y$ could be the water withdrawn from an aquifer and $x_i$ the capacity of the wells drilled by person $i$.

We assume $F' > 0$ and $F'' < 0$. These imply that $\exists k \geq 0 : \lim_{X \to \infty} F'(X) = k$, & $\lim_{X \to \infty} F''(X) = 0$. 

100
We assume further that person $i$’s output is

$$y_i = x_i \frac{F(X)}{X}$$

which means that she gets as her output a share of total equal to the share of inputs that she provides. Another way of thinking of this is that $F(X)/X$ is the average product of the input, and each agent gets the average product times the amount of input she provides. We can write this as

$$y_i = x_i \frac{F(x_i + x_{-i})}{x_i + x_{-i}}$$

where $x_{-i}$ is the vector of inputs provided by all agents other than $i$. If the cost of input is $p$ and the price of output is 1 then each agent seeks to maximize

$$\pi_i = x_i \frac{F(x_i + x_{-i})}{x_i + x_{-i}} - px_i$$

If she takes $X$ as given - as would make sense if there are many agents, each small with respect to the total, the FOCs are

$$\frac{F(X)}{X} = p$$

so that average product equals price.

Now look at the Pareto efficient outcome. We seek to maximize $F(X) - pX$ and the FOC is

$$\frac{\partial F}{\partial X} = p$$

or marginal product equals price. As $F(X)/X > F'(X)$, we see that the resource is over-used under a competitive regime. Note that under the assumptions specified on $F$, it is the case that

$$\lim_{X \to \infty} \left\{ F'(X) - \frac{F(X)}{X} \right\} = 0$$

so that with infinitely many agents the two outcomes are the same. They are also the same when there is only one agent.
23 Non-Convex Production Sets

If production sets are not convex, then

1. there may be no competitive equilibrium, as supply functions are not continuous

2. it is not the case that an Pareto efficient allocation can be supported as a CE

In this case we generally work with what is called a marginal cost pricing equilibrium, which is a form of regulated equilibrium.

**Definition 42.** A marginal cost pricing equilibrium (MCPE) is an allocation \((y^*_i, x^*_j)\) and a price vector \(p^*\) and wealth levels \(W_j\), \(\sum W_j = p^* \cdot \sum_j w_j + \sum_i p^* y^*_i\), such that

1. For each firm \(i\) the first order conditions for profit maximization are satisfied, i.e. \(p^* = \gamma_i \nabla F_i(y^*_i)\), for some \(\gamma_i > 0\).

2. For each consumer \(j\), \(x^*_j\) maximizes \(U_j(x_j)\) subject to \(p^* x_j \leq W_j\)

3. \(\sum_j x^*_j = \sum_j w_j + \sum_i y^*_i\)

So this is a competitive equilibrium except that firms are not necessarily maximizing profits, but they are satisfying the FOCs for profit maximization. Profits may be negative, so that they could be increased by closing down. So any CE is a MCPE, but the converse is not true.

Note that a MCPE satisfies the FOCs for PE: all marginal rates of substitution and transformation are equal. We can see that this is necessary for PE from proposition 41, which shows that a PE allocation is the maximum of a weighted sum of utilities subject to production and resource constraints.

Assume the sets \(Y_i\) can be described by functions \(F_i(y_i) \leq 0\). Then the maximization problem in proposition 41 is

\[
\begin{align*}
Max_{x_j, y_i} & \sum_j \alpha_j U_j(x_j), \quad \sum_j x_j = \sum w_j + \sum_i y_i, \quad F_i(y_i) = 0
\end{align*}
\]

The Lagrangian is

\[
L = \sum_j \alpha_j U_j(x_j) + \lambda \left\{ \sum_j x_j - \sum w_j - \sum_i y_i \right\} + \mu_i F_i(y_i)
\]
which gives as FOCs those of the MCPE.

In the US many regulated utilities are expected to price at or near marginal cost and then cover losses from the fixed elements of two-part tariffs.
Next we extend the static, deterministic model of competitive equilibrium considered above to a world of time and uncertainty. Assume that the world last for $T > 0$ time periods, indexed by $t = 1, ... , T$ and that any one of $S > 0$ states of the world may obtain, indexed by $s = 1, ... , S$. A state is a complete description of the evolution of the world over all time periods, a description so complete that it resolves all uncertainty. There are as usual $N$ distinct goods and services, indexed by $n = 1, ... , N$.

The major innovation is to distinguish commodities by the date and state in which they are available: $c_{jts}$ is now individual $j$’s consumption vector in period $t$ in state $s$. It is a vector in $\mathbb{R}^N$, and a complete description of agent $j$’s consumption is $\{c_{jts}\}_{t=1, \ldots, T; s=1, \ldots, S} \in \mathbb{R}^{NTS}$. We denote this by $c_j \in \mathbb{R}^{NTS}$. There are markets and prices for all $NTS$ commodities. So commodity $k$ at date $t$ and in state $s$ is a different commodity from $k$ at date $t'$ and in state $s'$. An umbrella if it rains is different from one if it is dry: $500,000$ when your house has just burned down is different from $500,000$ if your house is intact.

We have $TS$ times as many commodities, markets and prices as in the atemporal certain case, and the commodities are called “time-state-contingent commodities” or just “contingent commodities.” People have preferences over this extended commodity space $\mathbb{R}^{NTS}$, which will reflect their attitudes towards risk and time, and firms’ production sets $Y_i \subset \mathbb{R}^{NTS}$ contain production plans that are state-contingent and extend over time.

We can define a competitive equilibrium in this extended commodity space exactly as before. And as before it will be Pareto efficient, meaning that no alternative feasible allocation of goods between states, dates and people will make someone better off without making someone else worse off. Let $p_{jts}$ be the price of good $j$ in period $t$ and state $s$, $p \in \mathbb{R}^{NTS}$ the overall price vector, and $y_{its}$ be firm $i$’s production plan in period $t$ and state $s$, with $y_i \in \mathbb{R}^{NTS}$ its overall production plan. Firm $i$’s profit is as usual $p_i y_i$. So a competitive equilibrium in the state-contingent commodity world is defined as follows:

**Definition 43.** A competitive equilibrium of the economy with time-state-contingent commodities is a set of prices $p^* \in \mathbb{R}^{NTS}$, a set of production plans for each firm $y^*_i \in \mathbb{R}^{NTS}$ and a set of consumption plans for each
person $x_j^* \in \mathbb{R}^{NTS}$ such that

1. $y_i^*$ maximizes $\pi_i = p^*y$ for $y \in Y_i$

2. $x_j^*$ maximizes $U_j (x_j), p^*x_j = p^*w_j + \sum_i \theta_{ji} \pi_i = W_j$

3. $\sum_j x_j^* = \sum_j w_j + \sum_i y_i$

A consumer’s endowment $w_j$ is now in $\mathbb{R}^{NTS}$, giving endowments of goods by date and state. Note that except for the dimensionality of the commodity space this is exactly as the earlier definition of a competitive equilibrium, definition 39.

Trading occurs at the start of time, before a state of the world is realized. Agents trade and enter into contracts contingent on the state of the world (and time period), before the first time period and when the state is unknown.

One important point to understand is that in this framework firms’ profits are not stochastic variables, but certain numbers. They are known independently of the state of the world that obtains. The prices are known, and the firm produces and sells various state-contingent commodities. These are sold before the state is known. Once the state is known the firms deliver whatever they have contracted to deliver in that state, and have already been paid for this, and have already paid for any state-contingent inputs. So firms are fully hedged against uncertainty. Individuals’ consumption vectors are however uncertain: they are state-contingent. But conditional on a state and time period they are known with certainty.

The complexity of having state-contingent commodities and a commodity space of dimension $\mathbb{R}^{NTS}$ can be avoided by the device of securities, which are contracts that pay a specified amount (generally taken to be a unit of the currency, $\$1$) if and only if a particular state occurs. To focus on the role of securities assume now that there is only one time period but that there are many states, so that the commodity space is $\mathbb{R}^{NS}$, with each physical commodity being contingent on the state that occurs. Then a competitive equilibrium will exist and be efficient under the standard assumptions of earlier sections, but with $\mathbb{R}^{NS}$ prices and commodities.

As an alternative suppose that before uncertainty is realized, people and firms can trade securities that pay one dollar if and only if any particular state occurs, so that security $s$ pays $\$1$ if and only if state $s$ occurs, and securities are available for all states $s \in S$ (this assumption is often referred to as “complete insurance markets”). Once uncertainty is realized and the state is
known, agents can trade on risk-free spot markets (normal markets) using the money obtained from the securities they purchased that pay off in that state. Under certain conditions this structure can imitate the outcome that would occur if we had a competitive equilibrium with the full $NS$ markets. But instead we have only $N$ markets for goods and $S$ markets for securities, where $S$ is the number of states, and generally $N + S < NS$.

To see this, assume that $p^* \in R^{NS}$ (there is only one time period now) is the competitive equilibrium price vector with a full set of state-contingent commodity markets. A typical element is $p^*_{ks}$, the price of good $k$ in state $s$. Let $x^*_{jks}$ be the amount of good $k$ person $j$ bought in state $s$ at the competitive equilibrium, and let $y^*_{iks}$ be firm $i$’s supply of good $k$ in state $s$.

Now let $q_s$ be the price of a security that pays one unit if and only if state $s$ occurs. And let $\hat{p}_{ks}$ be the price of a unit of good $k$ on the spot market once uncertainty has been resolved and the state is known to be $s$. Then the cost of buying a unit of good $k$ if state $s$ occurs is $\hat{p}_{ks}q_s$. If $\hat{p}_{ks}q_s = y^*_{iks}$, then consumers and firms will make exactly the same choices as they made at the competitive equilibrium with the full set of contingent commodity markets. They will trade securities so as to fund these trades. To purchase $x^*_{jks}$ in this framework will cost agent $j$ $\hat{p}_{ks}q_s x^*_{jks}$ on the spot market once $s$ has been realized, so she will buy securities to the value of $r_{js}$ where

$$\sum_k \hat{p}_{ks}q_s x^*_{jks} = r_{js}$$

Her purchase of securities across all states $s$ must satisfy the budget constraint

$$\sum_s r_{js} = W_j$$

which follows from item 2 of definition 43.

So if consumers anticipate the prices that will rule in spot markets once uncertainty is resolved and believe these to satisfy $\hat{p}_{ks}q_s = p^*_{ks}$ $\forall s \in S, \forall k$ the outcome will be an efficient competitive equilibrium. The conclusion: $S$ securities markets and $N$ goods markets can replace $NS$ contingent commodity markets if agents have price expectations that mimic the prices that would have ruled on contingent commodity markets. (Arrow, The Role of Securities in an Optimal Allocation of Risk-Bearing, Review of Economic Studies 1964, first published in French in 1953.)