Rettevejledning

Each problem is weighed equally. Each sub-question of a given problem is weighted equally.

Problem 1 (25%)

a) A good answer should include:
   - Take a function $v = f\ (u)$, such that f is a strictly increasing real function.
   - Then, for each pair $x, x' \in \mathbb{R}_+^L$,
   $$v(x) \geq v(x') \Leftrightarrow\ u(x) \geq u(x') \Leftrightarrow x \succsim x'.$$
   meaning that also v represents preferences $\succsim$.
   - Informal arguments are fine.

   *

   Student may also go on to make the following contextual points that are not strictly needed:

   - The demand functions obtained from the utility function u or the utility function v are identical.
   - What matters is the preferences $\succsim$ and not the function chosen to represent such preferences.
   - In other words, the only information needed from utilities is ordinal: When comparing a pair of bundles $x, x' \in \mathbb{R}_+^L$, we only need to know the order of the alternatives, i.e. if $u\ (x) \geq u\ (x')$ or $u\ (x) \leq u\ (x')$ or $u\ (x) = u\ (x')$.

b) A good answer should include:
   - Let $v_i = f_i\ (u_i)$ and $v_j = f_j\ (u_j)$. Assume that society has Y dollars to divide, i.e. $x_i + x_j \leq Y$, and show that maximizing $W\ (x_i, x_j)$ leads to a different optimum than maximizing:
   $$\overline{W}(x_i, x_j) = v_i(x_i) + v_j(x_j).$$
   - Some may make a broader point on how to measure utility information, and to combine inter- and intra-personal comparisons.
   - Some may formulate the problem in the form of a Lagrangian or Fischer diagram to make their case, but informal arguments are fine.
   - Some may highlight appealing ways of distributing the Y dollars.

   *

   Student may also go on to make the following contextual points that are not strictly needed:

   - The above statement holds true even if $f_i = f_j$.
   - Bridging to c): $\overline{W}(x_i, x_j)$ gives always the same recommendation as $W(x_i, x_j)$ if and only if $f_i = a_i + bu_i$ and $f_j = a_j + bu_j$.

c) A good answer should include:
   - Utilitarianism builds on interpersonally comparable and cardinal information about utilities. u_i is equivalent to $v_i = f_i(u_i)$ only if $f_i(u_i) = a_i + bu_i$, i.e. u_i is invariant to any common positive affine transformation (or co-cardinal).

- Some may make a broader point on how to measure utility information, and to combine inter- and intra-personal comparisons.
- Some may formulate the problem in the form of a Lagrangian or Fischer diagram to make their case, but informal arguments are fine.

*

Student may also go on to make the following contextual point that is not strictly needed:

- Consumer theory builds on ordinal information about utilities. u_i is equivalent to $v_i = f_i(u_i)$ if f_i is a strictly increasing real function, i.e. u_i is invariant to any individual increasing transformation (or ordinal).

d) A good answer should include:
- Arrow (1951) showed that a set of compelling value judgements with ordinal information makes it impossible to derive a social welfare function.
- "If we exclude the possibility of interpersonal comparisons of utility, then the only methods of passing from individual tastes to social preferences which will be satisfactory and which will be defined for a wide range of sets of individual orderings are either imposed or dictatorial." (Arrow, 1963).

*

Student may also go on to make the following contextual point that is not strictly needed:

- Some may state the framework of the Arrow social welfare function/ collective choice rule and the associated axioms, to arrive at the Arrow's result:
  - Completeness of R.
  - Transitivity of R.
  - Unrestricted domain.
  - Weak Pareto principle (unanimity).
  - Nondictatorship.
  - Independence of irrelevant alternatives.

Problem 2 (25%)

a) A good answer should include:
- There is no need to define the maximin social welfare relation (SWR). Note that Strong Anonymity and Finite Anonymity are axioms of equal treatment. Weak Pareto and Strong Pareto axioms of sensitivity. (For definitions, see the other file: For problem 2.)
- Some informal reasoning along these lines:
  - The maximin SWR can be represented by a function:

$$W(x_1, x_2, \dots, x_t, \dots) = \inf_{t \geq 1} x_t$$

  - Strong Anonymity: yes (Any permutation works. Does not matter when the worst of comes, max their utility anyway.)
  - Finite Anonymity: yes (by satisfying Strong Anonymity.)

    - Weak Pareto: no (Strong Anonymity implies Finite Anonymity which contradicts Weak Pareto when SWR numerically representable.)
    - Strong Pareto: no (indifferent to making a well-off better off and keeping all other fixed.)
- Counterexample based on the streams in the exam: Indifferent between 1 and 2 (since the worst off has 1 in both cases), against Weak Pareto and Strong Pareto.

*

Students may use figures to highlight or define axioms. This is not strictly needed.

b) A good answer should include:
- There is no need to define the time-discounted utilitarian SWR. Note that Strong Anonymity and Finite Anonymity are axioms of equal treatment. Weak Pareto and Strong Pareto axioms of sensitivity. (For definitions, see the other file: For problem 2.)
- Some informal reasoning along these lines:
    - The time-discounted utilitarian SWR can be represented by a function:

$$W(x_1, x_2, \dots x_t, \dots) = (1-\beta)\sum_{t=1}^{\infty} \beta^{t-1} u(x_t)$$

    - Strong Pareto: yes (making someone better off and all other fixed gives larger total utility also with discounting.)
    - Weak Pareto: yes (by satisfying Strong Pareto.)
    - Finite anonymity: no (Strong Pareto implies Weak Pareto which contradicts Finite Anonymity when SWR numerically representable. Also, the rearranging of generations will matter, because we discount.
    - Strong Anonymity: no (by violating Finite Anonymity.)
- Counterexample based on the streams in the exam: Prefer 1 to 2, against Finite Anonymity and Strong Anonymity.

*

Students may use figures to highlight or define axioms. This is not strictly needed.

c) A good answer should include:
- Focus on the consequences of Finite Anonymity and Strong Pareto alone. These axioms have far-reaching implications in technological environments that satisfy a productivity condition.
- Productivity: If a feasible wellbeing stream is not non-decreasing (attachment 1, upper stream – but no figure needed). Then this wellbeing stream is feasible and inefficient (attachment 2, lower stream – but no figure needed). Hence, even this wellbeing stream is feasible (attachment 3, lower stream – but no figure needed).
- With this additional condition on technology, only non-decreasing streams are undominated. (Non-decreasing streams are sustainable.)

*

Students may make the contextual remark that for infinite wellbeing streams no SWR can satisfy both Strong Pareto and Strong Anonymity, or statements along the line:

- No complete SWR can satisfy both Weak Pareto and Finite Anonymity when some additional requirements are imposed.
- No continuous social welfare ordering (SWO) over infinite wellbeing streams can satisfy both Finite Anonymity and Weak Pareto.
- No explicitly describable SWO over infinite wellbeing streams can satisfy Finite Anonymity and Weak Pareto.
- No numerical representable SWO over infinite wellbeing streams can satisfy both Finite Anonymity and Weak Pareto.

Such statements are not strictly needed.

Problem 3 (25%)

a) A good answer should include:
   - The Lorenz curve relates the cumulative proportion of income units to the cumulative proportion of income received when units are arranged in ascending order of their income and takes the form of a straight line, the equality reference line, if and only if all units in the population receive the same income. If any units have unequal incomes the Lorenz curve is a convex function falling below the equality reference line (attachment 4).
   - More formally, but informal reasoning is fine:
     - A distribution of income for individuals (or households) can be represented by a continues (cumulative) distribution function F.
     - It is convenient to represent the observation with rank t by the inverse $F^{-1}(t)$ of the cumulative distribution function $F(x),\ 0 \leq x \leq \infty$, i.e. when F is continuous and $F(x) = t$ we have that $F^{-1}(t) = x$.
     - Assume that μ is the mean of F.

b) A good answer should include:
   - The Lorenz curve is a transformation of the cumulative distribution function, which satisfies the 1) Pigou-Dalton principles of transfers and 2) scale invariance.
   - Where 2) is the fact that inequality depends only on relative incomes, and not the levels of these incomes.
   - The Lorenz curve can be used as a ranking criterion of inequality when Lorenz curves do not intersect (then there is no ranking). Justified by Pigou-Dalton, which the Lorenz curve dominance criterion satisfies.
   - 1) It satisfies Pigou-Dalton because: A sequence of mean preserving transfers from rich to poor moves one from a lower to a higher Lorenz curve (attachment 5 – but no figure needed).

*

Students may go in more detail, but that is not needed.

- That 1) Pigou-Dalton is satisfied can be shown graphically (attachment 5).
- That 2) Scale invariance is satisfied can be shown mathematically:

$$X = \alpha Y \qquad X \sim G$$
$$G(X) = P(X \leq x) = P(\alpha Y \leq x) = F\left(\frac{x}{\alpha}\right)$$
$$G(x) = t \qquad x = G^{-1}(t)$$
$$F\left(\frac{x}{\alpha}\right) = t \qquad x = \alpha F^{-1}(t)$$
$$\mu_g = \alpha \mu_f$$
$$L_g(u) = \frac{1}{\mu_g}\int_0^u G^{-1}(t)dt = \frac{1}{\alpha\mu_f}\int_0^u \alpha F^{-1}(t)dt = \frac{1}{\mu_f}\int_0^u F^{-1}(t)dt = L_f(u)$$

c) A good answer should include:
- Now we are not only measuring shares, but are also able to say something about how large the differences in income are. This is basically dividing the Lorenz curve by the bracket at each point (attachments 6 and 7 -- for comparison). (Note that the M-curve, scaled conditional mean curves and normalized Lorenz curve are synonyms.)
- The attractive properties are:
  - For a fixed u, M(u) is the ratio between the mean income of the poorest 100u per cent of the population and the overall mean. Perfect equality if goes up and then to the right, perfect inequality if right and then up.
  - When a M-curve intersects the diagonal line once from above, the corresponding distribution exhibits lower inequality than a uniform (0, a) distribution below the intersection point and higher inequality than a uniform (0, a) distribution above the intersection point. (Uniform means that any income in an interval (0, a) occurs equally frequently.)
  - Visually sharper distinction between two different scaled conditional mean curves than two Lorenz curves. (Because of bound by the unit square.)

*

Student may also go on to make the following contextual point that is not strictly needed:

- Like the Lorenz curve, the normalized Lorenz curve satisfies Pigou-Dalton and scale invariance.

d) A good answer should include:
- They have different theoretical foundations (rely on different axiomatization), and that makes it difficult to judge their capacity as complementing measures of inequality.
- Of these measures, the Gini coefficient is the only one that can be explicitly expressed by the Lorenz curve.

e) A good answer should include:
- Definition of Gini's nuclear family:
  - Formally:
    $$C_k(F) = \int_0^1 u^k dM(u),\ k = 1,\ 2,\ \ldots$$
    But an informal discussion is satisfactory.
  - The entire set of moments of M, i.e. the family of dual inequality measures defined by C_k, uniquely determines the normalized Lorenz curve M. This

means that Gini's nuclear family satisfies the same properties as the normalized Lorenz curve.
    - Gini's nuclear family is C_k for $k = 1, 2, 3$.
- Reasons:
    - They have the same theoretical foundation (rely on different axiomatization).
    - Consistent with standard practice in the statistical literature for summarizing the information content of a distribution function the three first moments of the *M*-curve (C_1, C_2, C_3) can be used to summarize the information content of the *M*-curve and thus be used as primary quantities for measuring inequality.
    - Moreover, these three measures prove to supplement each other with regard to sensitivity to transfers at the lower, the central and the upper part of the income distribution. Thus, used together in applied work these three measures will provide information that will clarify whether changes (or differences) in inequality are due to changes in the upper, the central or the upper part of the income distributions.
    - C_1 is more sensitive to the lower part, C_2 to the central part, C_3 to the upper part. (C_2 is in fact the Gini coefficient.)

*

Student may also go on to make the following contextual points that are not strictly needed:

- The three first moments (C_1, C_2, C_3) are used to determine the location, the spread and the skewness and thus jointly provide essential information about the shape of the distribution function.
- Some may relate to the axiomatization of the dual approach and dual measures of inequality $\tilde{I}$, or the dual weight function $P(t)$.

Problem 4 (25%)

a) A good answer should include:
    - Data availability. Normal databases are typically restricted to post 1970/ 1980.
        - Top incomes are available for much longer horizons.
        - Top incomes can also be decomposed into sources of income.
        - Top incomes are fairly homogenous across countries, annual and long-run.
    - May be an important driver of overall inequality.
    - Top incomes and democracy. There might be concerns regarding the disconnect between the public and the concerns over growing influence of the wealthy.

b) A good answer should include:
    - Background on the evolution of top incomes that some may mention:
        - Over the last 30-40 years, there has been an increase in top incomes in many countries. This increase can largely be explained by increases in wages/ salaries (as opposed to capital incomes). While there has been an increase in many countries, this is not the case for all.

- Theoretical explanations:
    - Superstars. This is the most central force in the last 30-40 years. This is a winner-takes-it-all structure. The superstar argument goes as follows:
        - Those with the highest abilities are paid more.
        - This wage premium increases with their "reach".
        - Globalization and information technology has increased this reach. (Example of global force: International market for managers.)
    - Progressive taxation. This seems to be a major contributor. But there are large differences between countries. It is also not necessarily the most important factor. (Example: For Canada, the interactions with the US market is more important).
    - Wars and catastrophes. This is less relevant in the last 30-40 years, but might be mentioned by some as it has application is some cases. It argues that inequality has historically been reduces by warfare, revolutions, state collapse, and catastrophic plagues. This is associated with fall in top incomes through: Loss in capital incomes, equalization of earned income, changes in political regimes (for better or worse).

*

Student may also go on to make the following contextual points that are not strictly needed:

- Which countries the top income has increased (U curve since the early 1900: Drop in inequality, then increase again the last 30-40 years) and which countries it has been flat (L curve since the early 1900: Drop in inequality, then flattening out).
- Discussion of evolution in the first half of the last century.
- Discussion of other theories.

c) A good answer should include:

- Problem 1: Tax data typically only shows income that is taxed. Examples:
    - Some capital income might be tax exempt.
    - Capital gains only taxed when realized.
    - Deductions might have been subtracted.
    - Includes government transfers if these are taxed.
- Problem 2: Then there is the issue of tax avoidance and evasion. The incentive is strong among the rich to evade.
    - Some students may go in more detail on whether the rich or the poor evade taxes more based on data from a Swiss leaks from HSBC (tax haven), the Panama papers (tax havens) or Scandinavian tax amnesties. This research Focuses on rich Scandinavians, and it has established that the rich evade more than the poor do. Hence, the tax data underestimates the inequality in incomes.
- Yet, even though there are limitations to what we can say, the approach is not meaningless. Also, compared to household/ survey data, there is better coverage and longer time series.

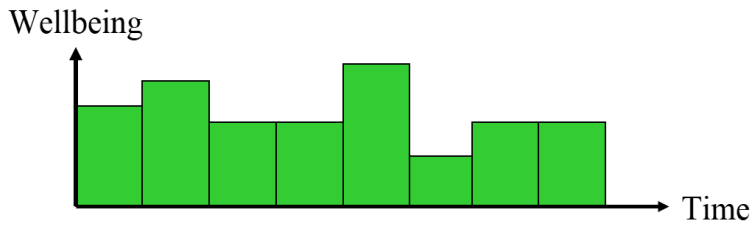
*

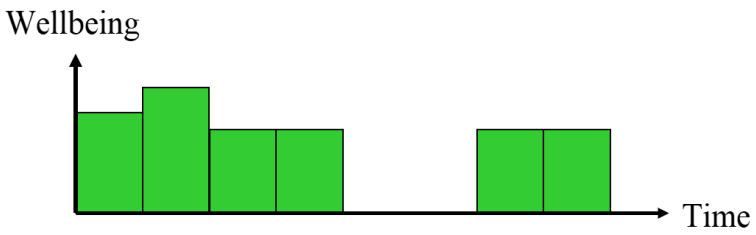
Student may also go on to make the following contextual points that are not strictly needed:

- Background: How are top incomes measured?
    - Top income share = Income of top percentile/ total income. Issues:
        - How many persons in one percentile?
        - Tabular data => Need to interpolate.
        - What is total income?
    - There are two approaches to get at total income: Either start with the income of those who filed taxes and estimate the income of those that did not (to get at the top income share, historically when not all filed taxed), or to use national accounts (and back total income out form a residual, like an accounting exercise).

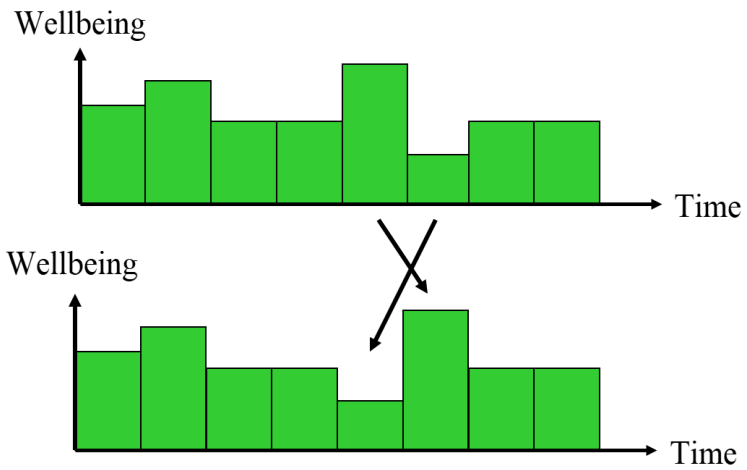
Attachments

Attachment 1

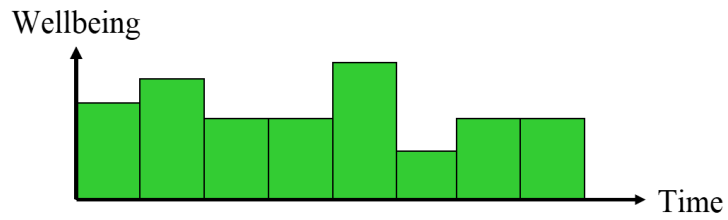


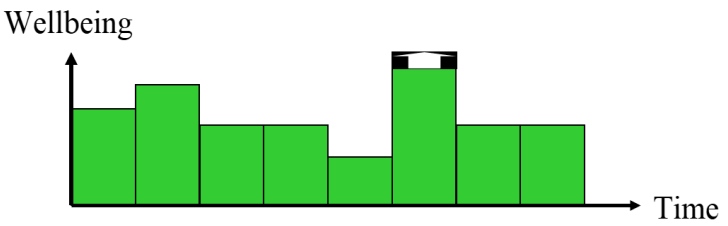


Attachment 2

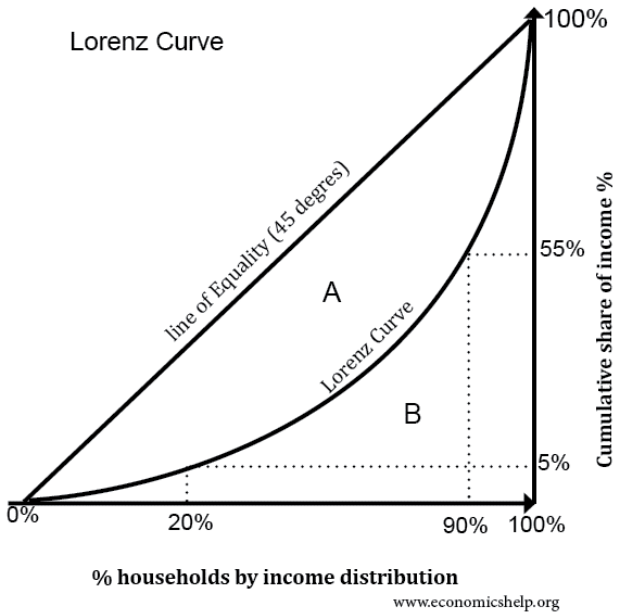


Attachment 3

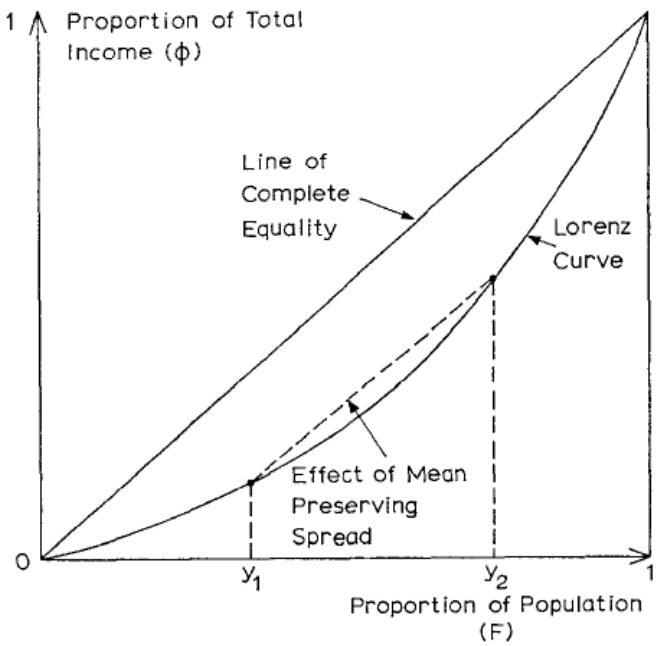




Attachment 4

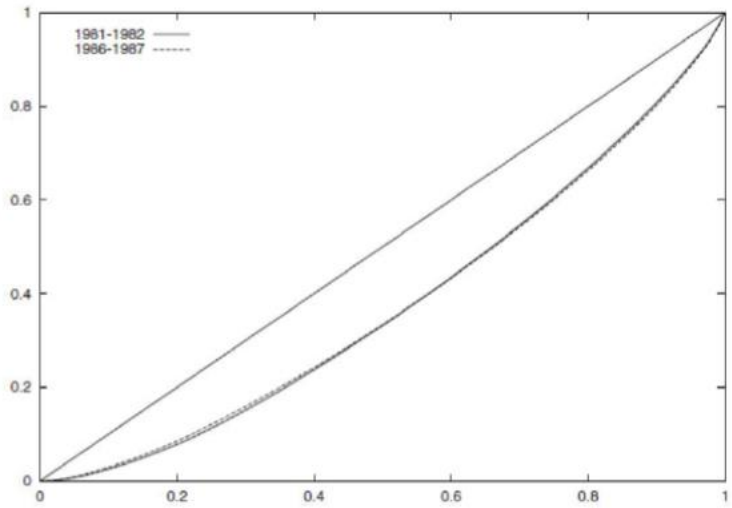


Attachment 5



Attachment 6

Figure 1. Lorenz curves for distributions of average annual earnings in Norway



Attachment 7

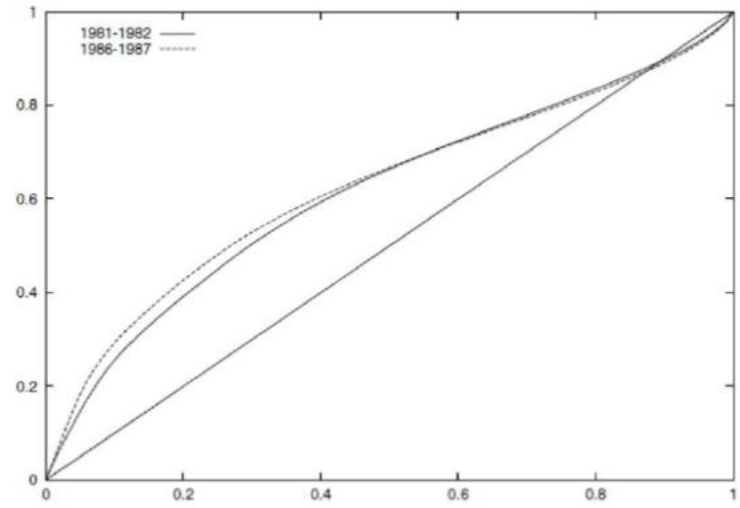



Figure 2. Normalized Lorenz curves for distributions of average annual earnings in Norway