Työpapereita
Working Papers

169

# DECOMPOSING THE GINI AND THE VARIATION COEFFICIENTS BY INCOME SOURCES AND INCOME RECIPIENTS

Ilpo Suoniemi

PALKANSAAJIEN TUTKIMUSLAITOS LABOUR INSTITUTE FOR ECONOMIC RESEARCH

October 2000

**Ilpo Suoniemi**
Labour Institute for Economic Research
Pitkänsillanranta 3A
FIN-00530 Helsinki
FINLAND
E-mail: ilpo.suoniemi@labour.fi

## Decomposing the Gini and the variation coefficients by income sources and income recipients

**ABSTRACT:** The Gini coefficient is not in general decomposable by population groups in terms of subgroup Ginis. On the other hand, there is a extensively used and well-founded decomposition of the Gini by income sources. In this paper a decomposition of the Gini by population groups is proposed which is simple and intuitively appealing. Here the decomposition by sources is utilized by defining a set of indicator functions for a partition of the population, and representing income as the sum of synthetic income sources. The approach is extended by treating each income source separately to give a general decomposition table. The table for the Gini is compared with a similar table obtained for the (square of) variation coefficient. The table elements give first order approximations to the change in the value of the inequality measure which is due to a proportional change in the income source affecting all individuals in the relevant group. Empirical applications of the method are illustrated by examples using Finnish household data.

**KEY WORDS:** Decompositions of inequality, Gini, variation coefficient

**TIIVISTELMÄ:** Gini-kerrointa ei yleisesti voida hajottaa väestöryhmien suhteen siten, että se lausutaan ryhmien Gini-kertoimien avulla. Sillä on kuitenkin paljon käytetty ja perusteltu hajotelma eri tulolajien suhteen. Tässä työssä esitellään yksinkertainen Gini-kertoimen hajotelma väestöryhmien suhteen, jolla on hyviä ominaisuuksia. Hajotelma käyttää hyväksi tulolajeittain tehtyä hajotelmaa. Tässä muodostetaan väestöositusta vastaava joukko osoitinmuuttujia ja esitetään tulot synteettisten tulolajien summana. Menetelmää laajennetaan käsittelemällä kukin tulolaji erikseen tällä tavalla, ja lopputuloksena saadaan yleinen hajotelmataulukko. Gini-kertoimen taulukkoa verrataan vastaavasti muodostettuun variaatiokertoimen (neliön) hajotelmataulukkoon. Taulukon alkiot antavat ensimmäisen kertaluvun arvion eriarvoisuusmittarin muutoksesta, kun kaikki väestöryhmään kuuluvat saavat saman suhteellisen lisäyksen tarkasteltavaan tulolajiin. Menetelmän sovellettavuutta esitellään Suomen kotitalousaineistosta tehtyjen laskelmien avulla.

**ASIASANAT:** Tulonjaon hajotelmat, Gini-kerroin, variaatiokerroin

# 1 Introduction

Decompositions of inequality measures offer useful methods of analysis by breaking down the temporal evolution of income inequality into more easily analysable components. The method can be used to assess the distributional role of factor income and the various items in the Government budget, see Atkinson (1997), and Atkinson et.al. (1995). Decompositions can be formed with respect to population subgroups and income sources, such as factor income, taxes and income transfers. Recently, Jenkins (1995) has used these methods to study the evolution of income inequality in the UK, in 1971-1986.

If population groups are considered, Shorrocks (1984) has convincingly argued that a natural summable decomposition can be obtained for only those inequality measures that belong to the family of *generalised entropy measures*, see also Cowell (1980). In the decomposition, the index is broken down into within- and between-group components. The latter is calculated using the group means and the former by using the within-group values of the measure. The square of the variation coefficient is a member in this family. It has an added advantage in having a natural decomposition in terms of income sources.

The Gini coefficient is the most extensively used summary measure of inequality. Lerman and Yitzhaki (1985) have presented a decomposition of the Gini coefficient as a weighted sum of the concentration coefficients of the income sources. Their and subsequent analysis has proved its usefulness in explaining inequality trends. In contrast, the Gini coefficient is not in general decomposable by population subgroups if one uses the group Ginis to calculate the within-group contributions. The Gini coefficient does not meet the conditions set by Shorrocks (1984) in the case of overlapping partitions of the income distribution. In this case a third component, a crossover effect, enters the calculations in addition to the within- and between-group contributions.

In the present paper it is argued that since the ranking of the individuals plays a central role in the computation of the Gini coefficient the aim to decompose the Gini in terms of subgroup Ginis runs counter to the intuition behind the Gini. A method of decomposing the Gini by population subgroups is proposed which is simple and intuitively appealing. It obeys the natural linearity of the Gini in terms of suitably defined concentration coefficients. To be more explicit, the decomposition by income sources is utilized here by defining a set of indicator functions for a given partition of the population. Multiplying the income variable by these indicators enables one to represent the income as a sum of synthetic income sources. The corresponding decomposition by these sources gives our method.

The approach can be extended by treating each income source separately to give a general decomposition table. In the table inequality is broken into elements that account simultaneously for both income sources and population subgroups. The rows in the table sum up to the total contribution of the population subgroup under consideration. Similarly, the column sums give the contributions of each income source to overall inequality. This decomposition of inequality is compact as it has no separate components for the within- and between-group contributions.

In this paper it is pointed out that the Gini coefficient and the square of the variation coefficient can be seen as members in a family of inequality measures that are defined as the mean of convex functions based on weighted differences in relative income. For comparison purposes a similar general decomposition table for the variation coefficient by population subgroups and income sources is presented. The tables are augmented by obtaining the sample variances of their individual elements to facilitate statistical inference.

The decomposition tables for the Gini and the variation coefficient are compared using empirical examples with Finnish data. Is is found out that the tables give same results qualitatively but the values in the table of the Gini are estimated more accurately. Comparing the differences in the individual elements across time periods and evaluating their statistical significance seem to give interesting insights into the evolution of inequality in the period of deep Finnish depression and subsequent economic recovery in the 1990's. The elements in the table reveal the marginal effects on total inequality which are due to changes in population group-specific income sources. A clear inequality-welfare ranking in recipient groups is revealed while budget neutral changes in social transfers are examined. In summary, it is argued that the simultaneous decomposition by income sources and recipients offers a useful tool for assessing the temporal evolution of income equality and the distributional role of the items in the Government budget.

The paper is organized as follows. Section 2 introduces the variation and Gini coefficients and reviews their main properties as inequality measures. In Section 3 the decomposition of the Gini by population subgroups is examined. The next two Sections set out and discuss the general decomposition tables for the Gini and the (square of) variation coefficients. Section 6 gives the sample variances of the elements in the decomposition table. Section 7 illustrates the method by empirical examples with Finnish household data. Detailed derivations of the sample statistics concerning the decomposition tables are given in the Appendix.

# 2 Properties of the variation and Gini coefficient

Shorrocks (1984) and Cowell (1980) have convincingly argued that a summable decomposition by population groups can be obtained for only those inequality measures that belong to the family of *generalised entropy measures*. Below our attention is confined to the square of the variation coefficient $I_2$ :

$$\begin{aligned} I_2 &= \frac{\sigma^2}{\mu^2} = \mathbf{E}\left(1 - \frac{y}{\mu}\right)^2 && (1) \\ &= \frac{2}{\mu^2}\int_0^\infty y(1-F)dy - 1. && (2) \end{aligned}$$

Variation coefficient allows inequality in variables that obtain zero values to be considered. This compares favorably with other members of the family, eg. the entropy measure $I_0$ where zero observations are ruled out by definition, Shorrocks (1984).

If one adopts Atkinson's (1970) position in defining inequality measures in terms of social welfare functions and inverts his argument, one obtains the following social welfare function corresponding to the $[0,1[$-normalised measure, $I_2/(1+I_2)$ :

$$W_{I_2} = \frac{1}{n}\sum \frac{y_i}{1+I_2}. \tag{3}$$

If the social welfare function is a representation of individual perceptions of social welfare, $W=(1/n)\sum U_i$, one obtains

$$\frac{\partial U_i}{\partial y_l} = \frac{\delta_{il}}{1+I_2} + \frac{2}{1+I_2}\left(1-\frac{y_l}{\bar{y}}\frac{1}{1+I_2}\right)\frac{y_i}{n\bar{y}}, \tag{4}$$

where $\delta$ denotes for the Kronecker delta, $\delta_{il}=1$, if $i=l$, and zero otherwise. By summation over $i$ one finds that the expression $\partial W/\partial y_l$ is not necessarily positive for large values of $y_l$. Therefore, the welfare function (3) may not be considered admissible if monotonicity in personal utility levels is taken as a necessary condition.

The decomposition of $I_2$ by population groups is given by

$$I_2 = \sum_i \pi_i \frac{\mu_i^2}{\mu^2} I_{2i} + \sum_i \pi_i \left(1-\frac{\mu_i}{\mu}\right)^2. \tag{5}$$

The first term is the within-group component of inequality, and the second term reflects inequality between population groups. The latter is calculated using the group means, $\mu_i$, and weights that are equal to population shares, $\pi_i$. In contrast, the within-group component is a weighted sum with weights which are formed as the product of the income shares and the relative (scaled with the population mean) means in the subgroups. In general, the weights do not add to one. For later use the equation (5) is written in a more compact form:

$$I_2 = \sum_i \pi_i \frac{\mu_i^2}{\mu^2}(I_{2i}+1) - 1. \tag{6}$$

Temporal changes in inequality can be traced to the changes in income across groups and in the structure of population by considering the index as a function of population shares and mean income in population groups. As an example we give the formula:[1]

$$\frac{\partial I_2}{\partial \log \mu_i} \quad = \quad -2\pi_i \frac{\mu_i}{\mu}\left((I_2+1) - \frac{\mu_i}{\mu}(I_{2i}+1)\right) \tag{7}$$

In the formula both the within- and the between-group contributions to inequality are involved. By (7) a proportional increase in income affecting population group $i$ increases inequality, if both $\mu_i \geq \mu$, and $I_{2i} > I_2$ hold. On the other hand, an increase in a (low

[1] The formula reflects a particular change in mean income where each member in the group gets the same proportional increase in income. Thereby, the value in the within-group inequality index is preserved. Note, $\partial \mu/\partial \log \mu_i = \pi_i \mu_i$.

income) group mean may decrease inequality, as measured by $I_2$, even if the within-group inequality is higher than the overall inequality.[2]

In addition, the measure $I_2$ has a simple decomposition by income sources. Let $y = \sum_1^m x_k$, then

$$Var(\sum_k x_k) = \sum_k Cov(x_k, y) \tag{8}$$
$$= \sum_k \rho_k \sigma_k \sigma, \tag{9}$$

where $\rho$ and $\sigma$ refer to the correlation coefficient and standard deviation, respectively. One obtains

$$I_2(y) = \sum \frac{\rho_k \sigma_k}{\sigma} I_2 = \sum \beta_k I_2 \tag{10}$$
$$= \sum \rho_k \frac{\mu_k}{\mu} \sqrt{I_{2k} I_2}. \tag{11}$$

The coefficients $\beta_k$ are easily calculated by regressing each income source, $x_k$, on total income $y$.

One may derive the partial derivative of the inequality index w.r.t. the mean of an income source to assess how inequality is affected by a proportional change in the income source $x_k$ that is equal across all individuals:

$$\frac{dI_2/I_2}{d\mu_k/\mu_k} = 2\left(\beta_k - \frac{\mu_k}{\mu}\right). \tag{12}$$

The overall inequality as measured by $I_2$ is decreased by increasing the mean of an income source if the regression coefficient is less than the relative share in total income.

Relative income distribution can be equivalently defined by its Lorenz curve, $LC_F$ :[3]

$$LC_F(p) = \frac{1}{\mu_y} \int_0^p F^{-1}(u) du \tag{13}$$
$$= \frac{1}{\mu_y} \int_0^{F^{-1}(p)} y dF_y \tag{14}$$
$$= \frac{1}{\mu_y} \mathbf{E} y \mathbf{1}(p), \tag{15}$$

where $F$ denotes for the cumulative distribution function, mean income is given by $\mu$, $\mathbf{E}$ refers to expectation, and $\mathbf{1}(p)$ is the indicator function for $\{y \leq F^{-1}(p)\}$. Lorenz curve

[2]In empirical analysis the above summable decomposition (5), say $I = \sum w_i x_i$, and the corresponding partial derivatives can be utilized to represent the temporal change in inequality, $dI = \sum \bar{w}_i dx_i + \sum \bar{x}_i dw_i$, where $\bar{w}_i = w_i + 0.5dw_i$. Jenkins (1995) utilizes this technique to get insight which factors are underlying the temporal evolution of inequality in the UK, in 1971-86.

[3]A point $(p, LC_F(p))$ on the curve tells the fraction of total income, $LC_F(p)$ that is earned by the least privileged p-percentage of the population. In the integral formulae of the paper we assume an absolutely continuous distribution with support $[0, \infty)$.

is a convex increasing curve which is defined in a unit square and lies below the diagonal line. If all have equal incomes the Lorenz curve lies on the diagonal line.

Similarly, the concentration curve of a variable $x$ w.r.t. to income $y$ is defined as

$$\begin{aligned} CC_F(p) &= \frac{1}{\mu_x}\int_0^{F^{-1}(p)} x dF_x & (16)\\ &= \frac{1}{\mu_x}\mathbf{E}x\mathbf{1}(p). & (17)\end{aligned}$$

The expected values $\mathbf{E}[y\mathbf{1}(p)]$, and $\mathbf{E}[x\mathbf{1}(p)]$ define the ordinates of the absolute Lorenz curve ($ALC$) and absolute concentration curve ($ACC$), respectively.

The Gini coefficient is traditionally defined as two times the area that is bounded by the Lorenz curve and the unit diagonal:

$$\begin{aligned} G(y) &= 2\int_0^1 (p - LC_F)dp & (18)\\ &= \frac{2}{\mu_y}\int_0^1 \Big(\mu_y - \mathbf{E}(y|y \le F^{-1}(p))\Big)\, p\ dp. & (19)\end{aligned}$$

Similarly, the absolute Lorenz curve ($ALC$) and absolute concentration curve ($ACC$), can be used to define the absolute Gini ($AG$), and concentration ($AC$) coefficients, respectively.

The Gini coefficient can be written in several alternative forms:

$$\begin{aligned} G(y) &= 1 - 2\int_0^1 LC_F(p)dp & (20)\\ &= 1 - \int_0^\infty (1-F)^2 dy & (21)\\ &= 1 - 2\mathbf{E}y(1-F) & (22)\\ &= \frac{1}{2}\mathbf{E}|y_1 - y_2|, & (23)\end{aligned}$$

where we have given the formulae in terms of the normalized variables, $\mathbf{E}y = 1$, with no loss of generality. The last mean-difference representation is the most useful one. Here $y_i, i = 1, 2$ refer to two independent copies of the distribution $F$. It is interesting to note that the square of the variation coefficient can be written in a similar form:[4]

$$I_2(y) = \frac{1}{2}\mathbf{E}|y_1 - y_2|^2. \tag{24}$$

To obtain a derivation of the Gini coefficient in terms of social welfare function (Atkinson 1970), $W = (1/n)\sum U_i$, one may define

[4] Both measures can be seen as members in a family of inequality measures where one considers pairwise (relative) income comparisons. These differences are weighted with a positive, convex and skew-symmetric function, say $\psi, \psi(-z) = \psi(z)$. These measures are Schur-concave functions, i.e. they obey the Pigou-Dalton condition.

$$U_k(y) = y_k \left(1 - \frac{1}{n}\sum_i \mathbf{1}(y_k \geq y_i)\right). \tag{25}$$

Here the marginal utility of individual income is decreased w.r.t. to the rank of incomes, ranging from one for the least privileged individual to zero for the most affluent.

Similarly as with the square of the variation coefficient the Gini coefficient $G(y)$, is decomposable by income sources. Let $y = \sum_1^m x_k$, then

$$G(y) = \sum_k \frac{\mu_k}{\mu_y} C(x_k, y), \tag{26}$$

where $C(x_k, y)$ is the concentration coefficient of the variable $x_k$ w.r.t. the income variable, $y$, Lerman & Yitzhaki (1985).

The above equation can be used to interpret how inequality is affected by a proportional change in the income source $x_k$ that is equal across all individuals:[5]

$$\frac{dG(y)/G(y)}{d\mu_k/\mu_k} = \frac{\mu_k C_k}{\mu_y G(y)} - \frac{\mu_k}{\mu_y}. \tag{27}$$

The elasticity formula for the Gini is in analogy with the formula for $I_2$ (12) if one interprets $AC_k/AG = \mu_k C_k/\mu_y G(y)$ as the Gini regression coefficient.

# 3 Decomposition of the Gini by income recipients

The decomposition of the Gini coefficient by income sources is used here to produce a simple decomposition by groups of income recipients which is in our opinion well-founded and potentially informative on the temporal changes in inequality.

Consider a partition of the population $\Omega = \sum A_i$, and the indicator variables $\mathbf{1}_i, i = 0, 1, \cdots, n$, $\mathbf{1}_i(a) = 1$ if $a \in A_i$, and zero otherwise. Writing $y = \sum_i \mathbf{1}_i y$, one obtains

$$G(y) = \sum_i \frac{\pi_i \mu_i}{\mu} C(y\mathbf{1}_i, y), \tag{28}$$

where $\mathbf{E}y\mathbf{1}_i = \pi_i \mu_i$, $\pi_i$ and $\mu_i$ stand for the population share and mean of the subgroup $i$, respectively.

The above decomposition of the Gini coefficient is canonical in the sense that it is based on the natural linearity of the expectation (or integral) operator, cf. $\mathbf{E}\sum X_i = \sum \mathbf{E}X_i$. The representation is a direct sum and has no separate components for within-group and between-group contributions. However, one may write

$$y = \sum_i (y - \mu_i)\mathbf{1}_i + \sum_i \mu_i \mathbf{1}_i. \tag{29}$$

[5]The examination is in analogy with the former partial derivatives of $I_2$. Here a supplementary condition is needed: The change is marginal in the sense that the original rank of observations w.r.t. the values of $y$ is unchanged. In particular, no ties in the values of $y$ are allowed. Thereby, the initial ranking can be maintained.

Multiplying both sides by the indicator $\mathbf{1}(p)$ for $y \leq F^{-1}(p)$, taking expectation and integrating, one obtains

$$
\begin{aligned}
G(y) &= \sum_i \frac{\pi_i \mu_i}{\mu} \left( \frac{2}{\pi_i \mu_i} \int_0^1 \mathbf{E}(\mu_i - y) \mathbf{1}_i(p) dp \right) \\
&+ \sum_i \frac{\pi_i \mu_i}{\mu} \frac{2}{\pi_i} \int_0^1 (\pi_i p - \mathbf{E} \mathbf{1}_i(p)) dp \\
&= \sum_i \frac{\pi_i \mu_i}{\mu} \left( \frac{2}{\pi_i \mu_i} \int_0^1 \int_0^p (\mu_i - y_u) \pi_i(u) du dp \right) \qquad (30) \\
&+ \sum_i \frac{\pi_i \mu_i}{\mu} \frac{2}{\pi_i} \int_0^1 \int_0^p (\pi_i - \pi_i(u)) du dp, \qquad (31)
\end{aligned}
$$

where $\mathbf{1}_i(p) = \mathbf{1}_i \mathbf{1}(p)$, $\pi_i(u)$ is the probability of inclusion in the group $i$ conditional on the income level, and $F(y_u) = u$. Note $\mathbf{E}\mathbf{1}_i(p) = \int_0^p \pi_i(u) du$, $\mathbf{E}_u \pi_i(u) = \pi_i$.

The first term (30) is composed as a weighted mean of the within-group terms that measure cumulative shortfall w.r.t. the group mean. The above within-group shortfall is generally not equal to the within-group Gini coefficient. Equality (up to a scaling constant) holds only if there are no overlaps in the distributions of the population groups considered. In other cases, the corresponding concentration curve of the synthetic income variable is horizontal in that part where no observations in the group in question are found if its support has say, two components. The second term (31) gives the corresponding between-group shortfall. In the Gini coefficient the weights are the income shares of the recipient groups, $\pi_i \mu_i / \mu$.

The above discussion reaffirms that in general the Gini coefficient is not decomposable in terms of within-group Ginis. This is in contrast with measures that are based on distance measures. For example, the Euclidean norm satisfies the Pythagorean Theorem which guarantees a canonical decomposition of the measure $I_2$. This may be considered as a handicap for the Gini. However, our 'decomposition of the Gini' explicitely accounts for the gaps in the group distribution while observations in other groups are encountered. A direct decomposition of the Gini in terms of the subgroup Ginis runs counter to the intuition behind the Lorenz curve. It is essential to account for the gaps in the distribution.[6]

If the within-group distributions are non-overlapping and have their supports with a single component one obtains a simple decomposition of the Gini coefficient in terms of the within-group Ginis:[7]

Consider population subgroups that have ordered, disjoint ranges in their income distributions, i.e. they can be ordered $i = 0, 1, \cdots, n$, such that $h \in A_i \Rightarrow y_h > y_j \forall y_j \in A_{i-1}$,

[6] If the group distributions are continuous and have a common support the above decomposition seems perfectly natural and the concentration coefficients take account of the relative intensities of the distributions.

[7] The result is well known and easily proved for our decomposition by induction.

then

$$G(y) = \sum_i \pi_i \frac{\pi_i \mu_i}{\mu} G_i(y) + B, \tag{32}$$

where $B$ is the Gini-coefficient that has been calculated using the subgroup means:[8]

$$B = 1 - 2 \sum_i \frac{\pi_i \mu_i}{\mu} \left( (1/2)\pi_i + \sum_{j>i} \pi_j \right). \tag{33}$$

If the formula (27) is expressed in absolute terms one finds that a revenue neutral, marginal change in the Government budget affecting income sources, k,l, $d\mu_k < 0, d\mu_k + d\mu_l = 0$, increases inequality as measured by the Gini if $C_l > C_k$, a result in line with Yitzhaki & Slemrod (1991). Above and in the following, welfare losses due to revenue neutral changes in tax and benefit schedules are ignored, or are assumed equal in the two groups under consideration.[9]

If this observation is applied to the income sources $\mathbf{1}_i y_k, \mathbf{1}_j y_k$ we obtain a useful Corollary: Let

$$C(\mathbf{1}_i y_k, y) > C(\mathbf{1}_j y_k, y), \tag{34}$$

then a revenue neutral (marginal) change, say in income transfers that benefits group $j$ at the expense of group $i$, decreases the overall Gini coefficient and increases social welfare as measured by (25).

# 4 Simultaneous decomposition by income sources and income recipients

The decompositions of the Gini coefficient by population groups and income sources can be combined to give a general, simultaneous decomposition. Consider, as above, the indicator variables $\mathbf{1}_i$, corresponding to a partition of the population, $i = 0, 1, \cdots, n$, and income sources $y = \sum_k y_k$, Writing $y = \sum_i \sum_k \mathbf{1}_i y_k$, one obtains

$$G(y) = \sum_i \sum_k \frac{\pi_i \mu_{ik}}{\mu} C(y_k \mathbf{1}_i, y) = \sum_i \sum_k \frac{1}{\mu} AC(y_k \mathbf{1}_i, y), \tag{35}$$

where $\mathbf{E} y_k \mathbf{1}_i = \pi_i \mu_{ik}$, where $\pi_i$ and $\mu_{ik}$ stand for the population share and mean of the income source $y_k$ in the subgroup $i$, respectively.

These components can be represented as a general decomposition in Table 1. Here the columns sum to the decomposition by income sources and the row sums correspond to the above decomposition of the Gini by population groups.

[8] See (22) with discrete random variables.

[9] Welfare losses that are due to price changes are of second order magnitude while the corresponding changes in social welfare are of the first order.

Table 1: Decomposition table for the Gini

| Group | Income source $1, \cdots, k, \cdots m$ | Group total |
|---|---|---|
| 0 | . | . |
| . | . | . |
| . | . | . |
| . | . | . |
| $i$ | $\cdots \quad \frac{\pi_i \mu_{ik}}{\mu} C(y_k \mathbf{1}_i, y) \quad \cdots$ | $\frac{\pi_i \mu_i}{\mu} C(y \mathbf{1}_i, y)$ |
| . | . | . |
| . | . | . |
| . | . | . |
| $n$ | . | . |
| Factor total | $\cdots \frac{\mu_k}{\mu} C(y_k, y) \cdots$ | $G(y)$ |

One is able to represent the variation coefficient $I_2$ in similar way by writing

$$y = \sum_i (y - \mu_i)\mathbf{1}_i + \sum_i (\mu_i - \mu)\mathbf{1}_i + \mu. \tag{36}$$

The elements in the first sum are mutually uncorrelated. Furthermore,

$$Cov\left((y - \mu_i)\mathbf{1}_i, (y - \mu_j)\mathbf{1}_j\right) = \pi_i Var(y_i)\delta_{ij} \tag{37}$$

$$Cov\left((y - \mu_i)\mathbf{1}_i, (\mu_j - \mu)\mathbf{1}_j\right) = 0 \tag{38}$$

$$\sum_j Cov\left((\mu_i - \mu)\mathbf{1}_i, (\mu_j - \mu)\mathbf{1}_j\right) = 0. \tag{39}$$

Therefore, one may write

$$\begin{aligned} Var(y) &= \sum_i Var\left((y - \mu_i)\mathbf{1}_i\right) + \sum_i \mathbf{E}(\mu_i - \mu)^2 \mathbf{1}_i \\ &= \sum_i \pi_i \sigma_i^2 + \sum_i \pi_i (\mu_i - \mu)^2. \end{aligned} \tag{40}$$

in self-explaining notation, and form a general decomposition for $I_2$ (Table 2) which is constructed in analogy with Table 1:

$$I_2 = \sum_{ik} v_{ik} \tag{41}$$

where

$$v_{ik} = \pi_i \frac{\mu_i^2}{\mu^2} \frac{\rho_{ik}\sigma_{ik}}{\sigma_i} I_{2i} + \frac{\pi_i \mu_{ik}(\mu_i - \mu)}{\mu^2}. \tag{42}$$

Table 2: Decomposition table for $I_2$

| Group | Income source $1, \cdots, k, \cdots m$ | Group total |
|---|---|---|
| 0 | . | . |
| . | . | . |
| . | . | . |
| . | . | . |
| $i$ | $\cdots \quad \pi_i \frac{\mu_i^2}{\mu^2} \beta_{ik} I_{i2} + \frac{\pi_i \mu_{ik}(\mu_i - \mu)}{\mu^2} \quad \cdots$ | $\pi_i \frac{\mu_i^2}{\mu^2} I_{i2} + \pi_i \frac{\mu_i}{\mu} \left( \frac{\mu_i}{\mu} - 1 \right)$ |
| . | . | . |
| . | . | . |
| . | . | . |
| $n$ | . | . |
| Factor total | $\cdots \beta_k I_2 \cdots$ | $I_2(y)$ |

The first part corresponds to the within-group component and summing over $k$ one gets the formula for total group income which corresponds to the first part in (5), see (9). In contrast, the second part for group income differs from the decomposition (5) by the terms $\pi_i(1 - \mu_i/\mu)$ which vanish in the summation over the groups. On the other hand, a simple calculation shows that the first part corresponds to $\beta_{iw} I_2$ where the coefficient $\beta_{iw}$ is calculated by regressing the synthetic (within-group) income source, $(y - \mu_i)\mathbf{1}_i$ on total income $y$, and the second part corresponds to $\beta_{ib} I_2$ where the coefficient $\beta_{ib}$ is calculated by regressing the synthetic (between-group) income source, $\mu_i \mathbf{1}_i$ on total income $y$, see (10). These two synthetic income sources are uncorrelated. Therefore, the sum of the beta's gives the coefficient $\beta_i$ of total income $y$ while regressing on the synthetic income source, $y_i \mathbf{1}_i$.

# 5 Additional comparisons

Above the main motivation for our decomposition of inequality by population groups has been to emphasize and track down the effects of a proportional income increase affecting all individuals in a given population group. By construction (top-down viewpoint) this effect can be followed further to however fine subdivision of a given population. In contrast, the summable decomposition by Shorrocks (1994) emphasizes monotonicity in the opposite direction. Monotonicity condition requires that if inequality is increased within a population group then total inequality in the population increases. Therefore, the latter (bottom-to-up viewpoint) approach is consistent to any further amalgamation of population groups (eg. up to the world population).

The (square of) variation coefficient has both decompositions available. Therefore, we can present an exact comparison with the standard decomposition. With no loss of generality let the population mean, $\mu = 1$. Consider the population, $\Omega$ with $N$ individuals and $n$ population groups $A_i, i = 1, \cdots, n$. The space of (relative) income distributions is isomorphic with $Y$, the positive cone of an N-dimensional Euclidean space. It has a orthogonal representation $Y = \bigoplus_i Y_i$ where the orthogonal subspaces $Y_i$ are isomorphic with the income distributions in the population subgroups under consideration. Let $y \in Y, y = \sum y_i, y_i \in Y_i, i = 1, \cdots, n$. By the Pythagorean Theorem

$$I_2(y) + 1 = (1/n)\|y\|^2 = (1/n)\sum_i \|y_i\|^2, \tag{43}$$

and $I_2(y)+1$ defines the same inequality ordering as $I_2(y)$. Let $e$, and $e_i = (1, 1, \cdots, 1)$ be the constant unit vectors in $Y$, and $Y_i$, respectively, and $\overline{\mu_i}$ the mean in subgroup $i, i = 1, \cdots, n$. The vectors $y_i - \overline{\mu_i}e_i$, and $\overline{\mu_i}e_i$ are orthogonal, and therefore

$$(1/n)\sum_i \|y_i\|^2 = (1/n)\sum_i \|y_i - \overline{\mu_i}e_i\|^2 + (1/n)\|\overline{\mu_i}e_i\|^2. \tag{44}$$

Our decomposition for the $I_2$ is exactly this equation with orthogonal components across population groups. The decomposition by Shorrocks (1994) makes an additional use of the Pythagorean Theorem to write

$$(1/n)\|\overline{\mu_i}e_i\|^2 = (1/n)\sum_i \|\overline{\mu_i}e_i - \overline{\mu}e\|^2 + (1/n)\|\overline{\mu}e\|^2. \tag{45}$$

in terms of the population mean $\mu$, recall $\mu = 1$. In this special case, the monotonicity condition can be seen as following directly from the Pythagorean Theorem.

# 6 Statistical analysis of inequality decomposition

The estimators of the inequality measures and the individual elements in the decomposition table are based on functions of sample moments. The asymptotic covariances of the sample moments are derived by their joint multinormal asymptotic distribution. The convergence results that are needed here are Kolmororov's Strong Law of Large Numbers and Central Limit Theorem and they hold under the standard conditions. The asymptotic variances for the estimators of the individual elements in the decomposition table are obtained by Rao's (1965) delta-method:

Let $f$ be a smooth function of parameters $\theta$. Suppose $\sqrt{n}(\overline{v} - \theta_0)$ has a limiting multinormal distribution with mean zero and the variance $\Omega$. Now $\sqrt{n}(f(\overline{v}) - f(\theta_0))$ has a limiting multinormal distribution with mean zero and the variance $V = J\Omega J^T$ where $J$ is the partial derivative matrix, $\partial f_i/\partial\theta_j$, evaluated at $\theta_0$.

In most cases the covariance matrix $\Omega = (\omega_{kh})$ is estimable by the relevant sample moments. Consider first the decomposition table for $I_2$. Write the sample equivalent,

$$\overline{v_{ik}} \quad = \quad \frac{1}{\bar{y}^2}\overline{\mathbf{1}_i x_k y} - \frac{1}{\bar{y}}\overline{\mathbf{1}_i x_k} = \frac{1}{\bar{y}^2}\frac{1}{n}\sum \mathbf{1}_i x_k y - \frac{1}{\bar{y}}\frac{1}{n}\sum \mathbf{1}_i x_k. \tag{46}$$

By the delta-method, $\sqrt{n}(\overline{v_{ik}} - v_{ik})$ has a limiting normal distribution with mean zero and the variance $V = J\Omega J^T$ where $\Omega$ is the estimated covariance matrix of the moments, $\overline{\mathbf{1}_i x_k y}$, $\overline{\mathbf{1}_i x_k}$, and $\overline{y}$.

The corresponding partial derivatives evaluated at the estimated values are given by

$$J = \left( \frac{1}{\bar{y}^2}, \frac{-1}{\bar{y}}, \frac{1}{\bar{y}^2} \left( \overline{\mathbf{1}_i x_k} - \frac{2}{\bar{y}} \overline{\mathbf{1}_i x_k y} \right) \right). \tag{47}$$

The terms in $\Omega$ are obtainable from[10]

$$\lim_n Cov\left(\sqrt{n}\ \overline{xz}, \sqrt{n}\ \overline{hy}\right) = Cov(xz, hy). \tag{48}$$

The analysis of the decomposition table for the Gini coefficient is similar but somewhat more complicated. First, write the absolute concentration coefficient in a more useful form. Define

$$H_x(y) = \int_0^y x(u) dF_y(u), \tag{49}$$

where $x(u)$ refers to the conditional expectation of $x$ given $\{y = u\}$. Note, $H_x(\infty) = \mathbf{E}x$, and $H_x(0) = 0$.

$$\begin{aligned}
AC(x, y) &= \mu_x - 2\int_0^1 \int_0^{F^{-1}(v)} x(y) dF(y) dv \\
&= \mu_x - 2\int_0^\infty x(y)(1 - F(y)) dF(y) \\
&= \int_0^\infty x(y)F(y)dF(y) - \int_0^\infty x(y)(1 - F(y))dF(y) \\
&= \int_0^\infty x(y)F(y)dF(y) + \int_0^\infty H_x(y) d(1 - F(y)) \\
&= \int_0^\infty x(y)F(y)dF(y) - \int_0^\infty \int_0^y x(u) dF_y(u) dF(y) \\
&= \int_0^\infty \int_0^y (x(y) - x(u)) dF_y(u) dF(y) \\
&= \int_0^\infty \int_0^\infty (x(y) - x(u)) \mathbf{1}\{y > u\} dF_y(u) dF(y). \qquad (50)
\end{aligned}$$

Therefore, the absolute concentration coefficient of the variable $x$ w.r.t. $y$ can be written as the mean value of the difference $(x(y_1) - x(y_2))\mathbf{1}\{y_1 > y_2\}$, where $y_1$, and $y_2$ are two independent copies from the distribution $F(y)$.[11]

[10] All covariance estimators can be given in small sample form, i.e. including the terms of order less than $1/n$. We refrain from it here since the covariance estimators are utilized in conjunction of the delta-method to obtain $\sqrt{n}-$ consistent estimators of the variance for a non-linear function of the estimators that the elements in $\Omega$ refer to. The remark holds also for the absolute concentration coefficient, see below.

[11] The mean-difference form for the absolute concentration coefficient may have some independent interest.

The sample estimate of the absolute concentration coefficient can be written as

$$\overline{AC}(x,y) = \frac{1}{n(n-1)} \sum_i \sum_{j \neq i} (x_i - x_j)\mathbf{1}_{ij}, \tag{51}$$

where $\mathbf{1}_{ij}$ is the indicator function for $\{y_i > y_j\}$. The estimator is unbiased. In fact it is minimum variance unbiased estimator of $AC$ in the class of all continuous distributions.

**Proof**: Consider an unbiased estimator $(x_1 - x_2)\mathbf{1}_{12}$. The vector of bivariate order statistics $\left((x_{(1)}, y_{(1)}), \cdots, (x_{(n)}, y_{(n)})\right)$, (ordering by the values of $y$) is a complete sufficient statistics in the class of all continuous distributions. By the Rao-Blackwell theorem:

$$\mathbf{E}\left((x_1 - x_2)\mathbf{1}_{12} | ((x_{(1)}, y_{(1)}), \cdots, (x_{(n)}, y_{(n)}))\right) \tag{52}$$

is the minimum variance unbiased estimator of $AC$. Furthermore, $(x_1 - x_2)\mathbf{1}_{12}(y_1, y_2)$ assumes the value of each pair $\left((x_{(i)}, y_{(i)}), (x_{(j)}, y_{(j)})\right)$ with probability $1/n(n-1)$. Therefore, the conditional expectation in (52) is equal to $\overline{AC}(x,y)$.

The variance of the estimator $\sqrt{n}\ \overline{AC}$ is given in Appendix A:

$$\lim Var\left(\sqrt{n}\ \overline{AC}\right) = \psi - 4AC^2, \tag{53}$$

where

$$\begin{aligned} \psi &= \int_0^\infty \left(\mu_x - x(y) + 2(x(y)F(y) - H_x(y))\right)^2 dF(y), \\ H_x(y) &= \int_0^y x(u)dF(y). \end{aligned}$$

By Kolmororov's Strong Law of Large Numbers the value of $\psi$ is consistently estimable by its sample equivalent if it exists.

To apply the delta-method one needs additionally (Appendix A)[12]

$$\lim Cov\left(\sqrt{n}\ \overline{z}, \sqrt{n}\ \overline{AC}(x,y)\right) \quad = \quad \phi - 2\mu_z AC, \tag{54}$$

where

$$\phi \quad = \quad \int_0^\infty z(y)\left(\mu_x - x(y) + 2(x(y)F(y) - H_x(y))\right) dF(y).$$

Finally, consider an individual element in the decomposition table:

$$\overline{v_{ik}} = \frac{\overline{AC}(\mathbf{1}_i x_k, y)}{\bar{y}}. \tag{55}$$

[12] For the decomposition table only the covariance with $\sqrt{n}\ \bar{y}$ is needed. Here we give a more general formula which can be applied to obtain the asymptotic variance for the corresponding concentration coefficient.

The variable $\sqrt{n}(\overline{v_{ik}} - v_{ik})$ has a limiting multinormal distribution with mean zero and the variance $V = J\Omega J^T$ where $\Omega$ is the covariance matrix of the estimators $\overline{AC}$ and $\overline{y}$. The partial derivatives that are calculated at their sample values are obtainable from

$$J = \left(\frac{1}{\bar{y}}, \frac{-\overline{AC}}{\bar{y}^2}\right). \tag{56}$$

The above formulae give in special cases the asymptotic sample variances of both the Gini- and the concentration coefficient. The distribution for the former has been derived earlier by Goldie (1977) and several others after him. Goldie examined the empirical cumulative distribution function and derived a functional convergence criterion. In contrast, the methods in Appendix A are elementary.

# 7 Empirical examples

The empirical applicability of the decompositions of the Gini and the variation coefficients is illustrated using Finnish data. First, the population subgroups to be examined are formed by partitioning the data with respect to the family type.[13] Inequality is examined using disposable household income.[14] The income sources that define Disposable Income are: Capital Income, Entrepreuneurial Income, Wage Income, Direct Taxes and Income Transfers.[15] Income Transfers are divided further, into two components, Social Benefits and Social Assistance.[16]

### Comparison of the Gini and variation coefficients

The decompositions of the Gini and variation coefficient are given in Tables 3 and 4. To simplify comparisons, the figures are shown in percentage points of the value of the inequality index. The decompositions show very similar results. If income sources are examined (bottom row) it is found out that Wage Income seems to have a substantial influence in increasing income inequality. On the other hand, Direct Taxes seem to have the most marked influence in the opposite direction. These effects are due to their high

[13]The groups are: Single adult, Childless Couple, Single Parent, Family with Children, Old Age household (all members over 65 years of age), and the residual group, Others (households larger than a single family).

[14]In calculating inequality each household member is assumed have access to an income level obtained by dividing total household income by an equivalence scale. The equivalent scale is here proportional to the square root of the number of household members, see Atkinson, et. al. (1995).

[15]Capital Income includes rents, dividends, interest payments and imputed rents from owner-occupied housing, Entrepreuneurial Income accrues from agriculture, forestry and firms. Wage Income consists of money wages, salaries and compensations in kind, deducting work expenses related to these earnings.

[16]Unemployment and sick benefits and occupational old age, disability and unemployment pensions that are related to previously earned income are included in the former component. Sosial Assistance includes transfers that are independent on past earnings (they may be currently means-tested), for example, housing benefits and child benefits, unemployment and welfare assistance and national old age, disability and unemployment pensions.

share in Disposable Income (Table 6 ). The Transfers in the category, Social Assistance are more effective in reducing inequality in comparison with those included in Social Benefits, see formula (27) and Tables 6 and 3.

The estimated values seem to indicate that the responses due to the income sources are represented clearly in the table for the Gini and there is no reason to prefer the alternative table. Similar observations hold if the population group contributions, the row sums are examined (the last column). The group-wise decomposition formula which is utilized here explicitely accounts for overlapping partitions of the income distribution. This may lie behind the pronounced observability which is present in Table 3.

The standard deviations of the elements in the tables indicate that the decomposition table of the Gini is estimated more accurately than the alternative table referring to the variation coefficient. For example, the value of the population Gini coefficient is estimated about six times as accurately as the square of variation coefficient.[17] Therefore, to obtain the same accuracy than in the Gini a larger sample size is needed in the case of the variation coefficient.

Recall the formulae:

$$G(y) = \frac{1}{2}E|y_1 - y_2| \tag{57}$$

$$I_2(y) = \frac{1}{2}E|y_1 - y_2|^2, \tag{58}$$

where $y_i, i = 1, 2$, refer to two independent copies of the (mean scaled) income variable. The absolute value function that enters the formula for the Gini is more robust to extreme observations than the square function used in defining the variation coefficient.[18] Generally, income distributions have heavy tails to the right. Therefore, contamination of the data is reflected more in the estimation of the variation coefficient.[19] More extensive empirical comparisons revealed that the above observations are quite general and do not overly depend on the partition of the data nor the observation period.[20]

### Marginal effects of income sources

The data in Tables 5 and 6 can be used to calculate the concentration coefficients of all synthetic income sources that are based on group-specific indicator variables. By (27) and Corollary (34) these values have a distinct role in assessing the instruments in reducing income inequality. Table 7 compares the concentration coefficients of population group specific income sources, Social Assistance and Direct Taxes. This source of

[17]However, note that square instead of the ordinary variation coefficient is used here. A straightforward application of the delta-method suggests that in relative terms, as in Tables 3 and 4, the variation coefficient is estimated twice as accurately as its square, i.e. three times less accurately than the Gini coefficient.

[18]To see this, consider the (generalized) derivatives of these functions.

[19]This observation is in analogy with the comparison between the median-regression (minimizing the sum of absolute deviations) and OLS regression.

[20]These methods have been applied to the Finnish Household Expenditure Survey data in several years starting from 1971 and using various partitions of the population.

information can be used to make some rough estimates how the population Gini changes when the tax-benefit schedules of the Government budget are adjusted and to find changes which are neutral in the Government budget and would decrease the overall Gini in the population, see Corollary (34).

As an example, the second column in Table 7 reveals that an increase in Social Assistance benefitting either the population group, 'Single Parent' or the group, 'Single adult', at the expense of the group, 'Pair with Children' which is neutral in the Government budget would decrease the overall Gini in the population. The last column shows that a similar change in Direct Taxes at the expense of the group 'Pair with Children' operates in the same direction in terms of the population Gini. Since tax and social benefit schedules can be tailored to take account of the household composition these observations may have some practical value.[21]

The decomposition of Gini coefficient can be compared with decompositions of the within-group Ginis (Tables 5 and 8, respectively). A cursorly examination reveals that our decomposition which includes and 'naturally' allocates both the between-group effects and the crossover effects that are due to overlapping partitions of the income distribution into their proper subgroups is much more informative. In addition, marginal increase in the income in the groups, Single adult, Single Parent and Old Age households would seem to decrease overall inequality.

Table 9 shows the residual effects w.r.t. each income source which are calculated as a difference between the last row in Table 5 and the column sums in Table 8. In the income source, Social Benefits the difference is largest, in relative terms. This is probably due to the fact that Income Transfers in this category accrue mainly to those in the low end of the income distribution. For example, within the group, Old Age households they seem to increase income inquality (Table 5) but if these households are considered with respect to their position in the overall income distribution we see hardly any effect on inequality (Table 8). Therefore, the within-group effects are relatively uninteresting. Furthermore, the above observations on inequality reducing changes in the Government budget cannot be made using the information in Table 8.

## Evolution of income inequality in the 1990's

Finally, we compare the decomposition tables of the Gini coefficient in 1990 and 1996. This gives a simple method to assess the temporal evolution of inequality in a most interesting epoch in the Finnish economy.[22] Table 11 gives a concise summary of the

[21] However, in this particular case one has to admit that the values are somewhat sensitive to our selection of the equivalence scale.

[22] In Finland, a long period of growth in the 1980's was abruptly ended by the deep Depression in the early 1990's. During 1990-1993 GDP dropped by 12 per cent in volume from the 1990 level. The unemployment rate rose from under 4 per cent to 16 per cent. Interestingly enough, the Depression left the relative inequality seemingly unaffected. A comparison of the decomposition tables in 1990 and 1994 confirms this. This has been mainly due to automatic stabilators operating in form of increased income transfers. In addition, unemployment rose more rapidly in the high wage, male dominated sector, manufacturing. However, the level of the public dept rose very rapidly, and this has been countered by

statistically significant differences in the elements of the Gini table between 1996 and 1990.[23] The value of the overall Gini has been increased. In addition, both Social Assistance and Direct Taxes have more influence in reducing income inequality in 1996 than in 1990. The effect of Social Benefits has moved to the opposite direction. Interestingly enough, there has been no significant change in the source, Wage Income. Change in Capital Income has been the main culprit in producing the clearly inequitable effects.[24]

At first sight Families with Children and Old Age households have been most adversely affected, if the effect of group income is considered. In addition one observes that the sign pattern within the columns, i.e. the income sources, is not uniform. For example, one may observe interesting differences in the effects of Direct Taxes (cf. the marginal effects of Social Assistance, above). However, the first observation should be treated with caution. The factoring is based on absolute concentration coefficients and the change in an individual element may be partioned as:

$$\Delta\left(\frac{\pi_i\mu_{ik}}{\mu}AC(y_k\mathbf{1}_i,y)\right) = \overline{\left(\frac{\pi_i\mu_{ik}}{\mu}\right)}\Delta AC(y_k\mathbf{1}_i,y) + \overline{AC(y_k\mathbf{1}_i,y)}\Delta\left(\frac{\pi_i\mu_{ik}}{\mu}\right) \quad (59)$$

where the expressions, say $\overline{x} = 0.5(x_1 + x_0)$ are calculated at the mid-point values.[25]

The temporal change in the decomposition table has been produced by dramatic changes in the income shares, i.e. the second element in (59). If population shares are considered one observes that the last, residual group 'Others' has radically lost its share, mainly at the expense of the group 'Pair with Children' (Table 14). Presumably, the Statistics of Finland has used more narrow definitions in forming the households in the 1996 sample in comparison with 1990. The following example is even more striking in illustrating the effects of a changing population structure to the value of the Gini coefficient.

### Education and income inequality

In the decade after 1979 the income inequality increased dramatically in the United Kingdom after a relatively stable period of three decades, Johnson (1996). Over the last two decades wage inequality and educational wage differentials have expanded markedly

raising the tax rates on wage income. In the period of economic recovery from 1994 up to the present, the rise in factor income, and particularly so in Capital Income, has been accompanied by cuts in real benefit levels. There has also been substantial cuts in the tax rates of capital income in the early 1990's. Simultaneously with a high GDP growth, particularly affecting the export manufacturing industries and the asset values, the unemployment rate has been left at a relatively high level (14 per cent in 1996).

[23]The numerical values of these changes are important in their own right. There has been a relatively large change in the group, Single Parent, but small number of observations involved seems to take its toll while statistical significance is assessed. Here space considerations deny a more extensive discussion of empirical results.

[24]The factor share of labour is today at a relatively low level (about 55 per cent in 1996 while it has been over 60 per cent in the 1970's and 1980's).

[25]More detailed examinations of the inequality change may use a further partioning of the term. $\Delta\left(\frac{\pi_i\mu_{ik}}{\mu}\right) = \overline{\left(\frac{\mu_{ik}}{\mu}\right)}\Delta\pi_i + \overline{\pi_i}\Delta\left(\frac{\mu_{ik}}{\mu}\right)$.

in the USA. There are several explanations for this development. A popular one involves a positive productivity shock that reward high skills relatively more than in the past. These skills are acquired by high education levels and increased use of new information technology is frequently seen as a supplementary factor in the skill-biased productivity change. Some people see globalization of world trade as an additional factor that adversely affects low-skill workers in the developed economies through competition from emerging economies.

If the composition of the labour force and labour supply are held relatively constant, the changes in the demand for labour are transmitted by either higher wages or higher employment rates for high-skill workers. In both cases high-skill workers increase their share in wage income. This explanation has been extensively discussed and challenged by Atkinson (1999) and (2000).

Below Finnish data is examined to see whether such development has had any marked effect on the recent income distribution. The population subgroups are formed by partitioning the data with respect to education level attained by the household head. The classification is based on the Unesco International Standard Classification of Education (ISCED, 1974). Here five broad education levels are used.[26]

For comparison, look first at a stable period in Finnish income inequality. Table 12 gives differences in the tables of the Gini coefficient between 1976 and 1990. The Gini coefficient of disposable income has remained practically constant in this time period, the change is a mere 0.2 percentage points which is statistically insignificant. On the other hand, one sees remarkable changes in the column of Wage Income. However, the changes in absolute concentration coefficients are mainly due to the changes in the respective income shares (Table 13). In addition, the income shares closely follow the corresponding change in population shares (Table 14). The pressure on income distribution which is seeemingly due to wages in Table 12 has been produced by the higher education levels obtained by Finnish workers in 1990 as compared to 1976. Change in the composition of the population not in the educational wage differentials has been the underlying factor for the change. Interestingly enough the composition change is balanced by other income sources and does not show up in the values of the Gini coefficient.

Table 15 gives the corresponding differences in the tables of the Gini coefficient between 1990 and 1996. In this more recent case one observes no clear pattern in the column for Wage Income. On the other hand, the changes in income shares are again closely followed by the respective changes in population shares (Tables 16 and 14).[27]

Expansion of the educational wage or employment differentials has found little support in the Finnish experience even though the economy has experienced mass unemployment and dramatic restructuring of the economy in the 1990's. However, developments in the

[26]The groups are: Basic education (not exceeding 9 years), Lower level of upper secondary education (10-11 years), Upper level of upper secondary education (12 years), Lower levels of tertiary education (13-15 years), and Higher degree of tertiary education (at least 16 years of education).

[27]The observed changes in the contribution of Wage Income to total inequality in Disposable Income are so small that the presentation of their standard errors serves no particular merit in the present context.

wage differentials are clearly important and need to watched in the future. Presently, developments in Capital Income are the main source for the increase in relative income inequality if the sources of factor income are under consideration.

# References

Atkinson, A.B. (1970), "On the measurement of inequality", *Journal of Economic Theory*, **2**, 244-263.

Atkinson, A.B. (1997), "Bringing income distribution in From the Cold", *The Economic Journal*, **107**, 297-321.

Atkinson, A.B. (1999), "Is rising income inequality inevitable? A critique of the transatlantic concensus", *1999 Wider Annual Lecture*.

Atkinson, A.B. (2000), "The changing distribution of income: Evidence and explanations", *German Economic Review*, **1**, 3-18.

Atkinson, A.B., Rainwater, L. & T.M. Smeeding (1995), "Income Distribution in OECD countries: Evidence from the Luxembourg income study." *OECD*, Paris.

Cowell, F.A. (1980), "On the structure of additive inequality measures", *Review of Economic Studies*, 47, 521-531.

Goldie, C.M. (1977), "Convergence theorems for empirical Lorenz curves and their inverses", *Advances in Applied Probability*, 9, 765-791.

Jenkins, S.P. (1995), "Accounting for inequality trends: Decomposition analyses for the UK, 1971-86", *Economica*, **62**, 29-63.

Johnson, P. (1996), "The assessment: Inequality", *Oxfors Review of Economic Policy*, **12**, Spring, 1-14.

Lerman, R.I. & S. Yitzhaki (1985), "Income inequality effects by income source: A new approach and applications to the United States", *The Review of Economics and Statistics*, **67**, 151-156.

Shorrocks, A.F. (1984), "Inequality decomposition by population subgroups", *Econometrica* **52**, 1369-1385.

Yitzhaki, S., & J. Slemrod (1991), "Welfare dominance: An application to commodity taxation", *American Economic Review*, **81**, 480-496.

# Appendix A

In the following, one considers a sample of independent observations. Start with the covariance:

$$
\begin{aligned}
& Cov\left(\sqrt{n}\ \overline{AC}(z,y),\ \ \sqrt{n}\ \overline{AC}(x,y)\right) \\
= \ & Cov\left(\frac{1}{n-1}\sum_{j\neq i}(z_i - z_j)\mathbf{1}_{ij}\ ,\ \frac{1}{n-1}\sum_{k}\sum_{l}(x_k - x_l)\mathbf{1}_{kl}\right) && (60) \\
= \ & \frac{1}{n-1} Cov\left((z_i - z_j)\mathbf{1}_{ij}, \sum_{k\neq i,j}(x_i - x_k)\mathbf{1}_{ik} + \sum_{k\neq i,j}(x_k - x_i)\mathbf{1}_{ki}\right. \\
+ \ & \left.\sum_{k\neq i,j}(x_k - x_j)\mathbf{1}_{kj} + \sum_{k\neq i,j}(x_j - x_k)\mathbf{1}_{jk} + (x_i - x_j)\mathbf{1}_{ij} + (x_j - x_i)\mathbf{1}_{ji}\right). && (61)
\end{aligned}
$$

To simplify the above formula one has to evaluate:[28]

$$
\begin{aligned}
\psi_{zx} = \ & \int_0^\infty \int_0^\infty \int_0^\infty (z(y) - z(u))(x(y) - x(v))\mathbf{1}\{y > u\}\mathbf{1}\{y > v\} dF_y dF_u dF_v \\
+ \ & \int_0^\infty \int_0^\infty \int_0^\infty (z(y) - z(u))(x(v) - x(y))\mathbf{1}\{y > u\}\mathbf{1}\{v > y\} dF_y dF_u dF_v \\
+ \ & \int_0^\infty \int_0^\infty \int_0^\infty (z(u) - z(y))(x(v) - x(y))\mathbf{1}\{u > y\}\mathbf{1}\{v > y\} dF_y dF_u dF_v \\
+ \ & \int_0^\infty \int_0^\infty \int_0^\infty (z(u) - z(y))(x(y) - x(u))\mathbf{1}\{u > y\}\mathbf{1}\{y > v\} dF_y dF_u dF_v \\
= \ & \int_0^\infty \int_0^y \int_0^y (z(y) - z(u))(x(y) - x(v)) dF_u dF_v dF_y && (62) \\
+ \ & \int_0^\infty \int_y^\infty \int_0^y (z(y) - z(u))(x(v) - x(y)) dF_u dF_v dF_y && (63) \\
+ \ & \int_0^\infty \int_y^\infty \int_y^\infty (z(u) - z(y))(x(v) - x(y)) dF_u dF_y dF_v && (64) \\
+ \ & \int_0^\infty \int_0^y \int_y^\infty (z(u) - z(y))(x(y) - x(v)) dF_u dF_v dF_y && (65) \\
= \ & \int_0^\infty (\mu_x - x(y) + 2(x(y)F(y) - H_x(y)))\,(\mu_z - z(y) + 2(z(y)F(y) - H_z(y)))\, dF_y, && (66)
\end{aligned}
$$

above

$$
\begin{aligned}
H_x(y) &= \int_0^y x(u) dF_u, \\
H_z(y) &= \int_0^y z(u) dF_u.
\end{aligned}
$$

[28] In principle the expected values in the formula of $\psi$ could be calculated by using the corresponding sample moments. In practice, the amount of calculation involved is overwhelming.

The final equation (66) follows by partial integration. For example, examine (62):

$$\begin{aligned}
&\int_0^\infty\int_0^y\int_0^y (z(y)-z(u))(x(y)-x(v))dF_u dF_v dF_y \\
=\ &\int_0^\infty\int_0^y (z(y)F(y)-H_z(y))(x(y)-x(v))dF_v dF_y \\
=\ &\int_0^\infty (z(y)F(y)-H_z(y))(x(y)F(y)-H_x(y))dF_y. \qquad (67)
\end{aligned}$$

Similarly,

$$\begin{aligned}
&\int_0^\infty\int_y^\infty\int_0^y (z(y)-z(u))(x(v)-x(y))dF_u dF_v dF_y \\
=\ &\int_0^\infty (z(y)F(y)-H_z(y))(\mu_x-H_x(y)-x(y)(1-F(y)))dF_y, \qquad (68) \\
&\int_0^\infty\int_y^\infty\int_y^\infty (z(u)-z(y))(x(v)-x(y))dF_u dF_y dF_v \\
=\ &\int_0^\infty (\mu_z-H_z(y)-z(y)(1-F(y)))(\mu_x-H_x(y)-x(y)(1-F(y)))dF_y, \qquad (69) \\
&\int_0^\infty\int_0^y\int_y^\infty (z(u)-z(y))(x(y)-x(v))dF_u dF_v dF_y \\
=\ &\int_0^\infty (\mu_z-H_z(y)-z(y)(1-F(y)))(x(y)F(y)-H_x(y))dF_y. \qquad (70)
\end{aligned}$$

By collecting the various parts one gets $\psi_{zx}$. By formula (61):[29]

$$\lim Cov\left(\sqrt{n}\ \overline{AC}(z,y), \sqrt{n}\ \overline{AC}(x,y)\right) = \psi_{zx} - 4AC(z,y)AC(x,y). \qquad (71)$$

To apply the delta-method, one needs to calculate[30]

$$Cov\left(\sqrt{n}\ \bar{z}, \sqrt{n}\ \overline{AC}(x,y)\right) = Cov\left(z_k, \frac{1}{n-1}\sum_i\sum_{j\neq i}(x_i-x_j)\mathbf{1}_{ij}\right). \qquad (72)$$

Dropping the terms that are of order less than $n$, one obtains

$$\begin{aligned}
Cov\left(\sqrt{n}\ \overline{z}, \sqrt{n}\ \overline{AC}(x,y)\right) &= \frac{1}{n-1}Cov\left(z_k, \sum_{j\neq k}(x_k-x_j)\mathbf{1}_{kj}\right) \\
&+ \frac{1}{n-1}Cov\left(z_k, \sum_{i\neq k}(x_i-x_k)\mathbf{1}_{ik}\right). \qquad (73)
\end{aligned}$$

[29] It is found out that equation (71) directly proves the consistency of our estimator of the absolute concentration coefficient in the quadratic mean, provided that $\psi$ is finite. Below, we will need asymptotic normality, i.e. bounded higher order moments to guarantee the Central Limit Theorem to hold. Conditions of the Dominated Convergence Theorem may be invoked for these purposes. In addition, the value of $\psi$ can be estimated consistently by the corresponding sample moment.

[30] In fact, for our decompositon we need only covariance w.r.t $\sqrt{n}\ \bar{y}$. Here we present a more general formula which can be utilized to calculate the sampling variance of the Gini- and concentration coefficients, for example see Table 7.

To estimate the above formula one must evaluate

$$\begin{aligned}
\phi &= \int_0^\infty \int_0^\infty z(y)(x(y)-x(u))\mathbf{1}\{y>u\}dF_y dF_u \\
&+ \int_0^\infty \int_0^\infty z(y)(x(u)-x(y))\mathbf{1}\{u>y\}dF_y dF_u \qquad (74)\\
&= \int_0^y z(y)\left(\int_0^y (x(y)-x(u))dF_u - \int_y^\infty (x(y)-x(u))dF_u\right) dF_y \\
&= \int_0^\infty z(y)\left(\mu_x - x(y) + 2(x(y)F(y) - H_x(y))\right) dF_y. \qquad (75)
\end{aligned}$$

One obtains

$$\lim Cov\left(\sqrt{n}\,\overline{z}, \sqrt{n}\,\overline{AC}(x,y)\right) = \phi - 2\mu_z AC. \tag{76}$$

Returning to a general element in the decomposition table:

$$\overline{v_{ik}} = \frac{\overline{AC}(\mathbf{1}_i x_k, y)}{\bar{y}}, \tag{77}$$

and $\sqrt{n}(\overline{v_{ik}} - v_{ik})$ has a limiting normal distribution with mean zero and the variance $V = J\Omega J^T$ where $\Omega$ is the estimated covariance matrix of $\overline{AC}$, and $\overline{y}$.

The corresponding partial derivatives evaluated at the estimated values are given by

$$J = \left(\frac{1}{\bar{y}}, \frac{-\overline{AC}}{\bar{y}^2}\right). \tag{78}$$

Finally, we note that in the case of sampling weights that are independent of the variables under examination, one substitutes

$$\overline{AC}(x,y,h) = \frac{1}{n(n-1)}\sum_i \sum_{j\neq i} h_i h_j (x_i - x_j)\mathbf{1}_{ij}, \tag{79}$$

for the estimators

$$\overline{AC}(x,y) = \frac{1}{n(n-1)}\sum_i \sum_{j\neq i} (x_i - x_j)\mathbf{1}_{ij}. \tag{80}$$

Above the sum of the sample weights $h_i$ is equal to $n$.

Table 3: Decomposition of the normalized Gini coefficient in 1996

| Sub-Group | Capital Income | Entrepr. Income | Wage Income | Social Assistance | Social Benefits | Direct Taxes | Disposable Income |
|---|---|---|---|---|---|---|---|
| Single adult | 0.64 | -0.04 | 6.23 | -4.97 | -2.53 | -1.98 | -2.64 |
| | 0.52 | 0.39 | 2.18 | 0.74 | 0.56 | 0.94 | 1.88 |
| Pair, no child | 4.56 | 5.02 | 39.79 | -0.22 | 4.85 | -19.43 | 34.57 |
| | 1.15 | 1.53 | 4.10 | 1.06 | 1.29 | 2.09 | 3.67 |
| Single Parent | 0.36 | -0.11 | -0.62 | -3.51 | 0.41 | -0.31 | -3.78 |
| | 0.39 | 0.13 | 1.43 | 0.62 | 0.27 | 0.76 | 1.36 |
| Pair, children | 10.62 | 11.08 | 67.84 | -4.78 | 1.22 | -31.95 | 54.03 |
| | 1.87 | 2.44 | 4.87 | 1.06 | 0.53 | 2.31 | 4.21 |
| Old Age | 0.58 | 0.11 | 0.12 | -3.43 | 0.41 | -2.10 | -4.31 |
| | 0.79 | 0.21 | 0.12 | 0.65 | 1.96 | 1.03 | 2.05 |
| Other | 5.91 | 4.33 | 14.10 | 1.21 | 7.06 | -10.48 | 22.13 |
| | 2.46 | 1.54 | 2.28 | 0.71 | 1.88 | 1.80 | 3.77 |
| All | 22.67 | 20.40 | 127.47 | -15.71 | 11.43 | -66.25 | 100.00 |
| | 3.18 | 3.12 | 4.73 | 1.48 | 2.93 | 2.10 | 2.70 |

To simplify the comparisons the estimated values (odd rows) and standard deviations (even rows) have been expressed as percentage points in the overall inequality.

Table 4: Decomposition of the normalized variation coefficient in 1996

| Sub-Group | Capital Income | Entrepr. Income | Wage Income | Social Assistance | Social Benefits | Direct Taxes | Disposable Income |
|---|---|---|---|---|---|---|---|
| Single adult | 0.43 | 0.29 | 3.18 | -2.97 | -1.53 | -1.08 | -1.68 |
| | 0.40 | 0.50 | 1.75 | 0.39 | 0.29 | 0.78 | 1.60 |
| Pair, no child | 6.63 | 8.02 | 30.03 | 0.10 | 2.29 | -17.68 | 29.39 |
| | 3.42 | 3.69 | 4.64 | 1.13 | 0.82 | 3.75 | 6.24 |
| Single Parent | 0.41 | -0.06 | 0.28 | -1.98 | 0.25 | -0.56 | -1.66 |
| | 0.48 | 0.07 | 1.63 | 0.40 | 0.24 | 0.92 | 1.37 |
| Pair, children | 10.48 | 14.16 | 38.61 | -3.47 | 0.45 | -21.08 | 39.14 |
| | 3.17 | 6.59 | 5.95 | 1.12 | 0.33 | 3.12 | 8.37 |
| Old Age | 1.69 | 0.21 | 0.17 | -2.16 | 2.72 | -3.04 | -0.39 |
| | 1.32 | 0.20 | 0.16 | 0.39 | 2.41 | 1.47 | 2.47 |
| Other | 20.70 | 5.83 | 8.87 | 1.67 | 14.97 | -16.84 | 35.20 |
| | 17.09 | 2.21 | 1.65 | 0.69 | 9.94 | 8.80 | 18.34 |
| All | 40.34 | 28.45 | 81.14 | -8.81 | 19.15 | -60.27 | 100.00 |
| | 17.04 | 7.58 | 9.02 | 1.96 | 9.78 | 8.17 | 16.89 |

To simplify the comparisons the estimated values (odd rows) and standard deviations (even rows) have been expressed as percentage points in the overall inequality.

Table 5: Decomposition of the Gini coefficient in 1996

| Sub-Group | Capital Income | Entrepr. Income | Wage Income | Social Assistance | Social Benefits | Direct Taxes | Disposable Income |
|---|---|---|---|---|---|---|---|
| Single adult | 0.15 | -0.01 | 1.42 | -1.13 | -0.58 | -0.45 | -0.60 |
| Pair, no child | 1.04 | 1.14 | 9.08 | -0.05 | 1.11 | -4.43 | 7.89 |
| Single Parent | 0.08 | -0.02 | -0.14 | -0.80 | 0.09 | -0.07 | -0.86 |
| Pair, children | 2.42 | 2.53 | 15.47 | -1.09 | 0.28 | -7.29 | 12.32 |
| Old Age | 0.13 | 0.02 | 0.03 | -0.78 | 0.09 | -0.48 | -0.98 |
| Other | 1.35 | 0.99 | 3.22 | 0.28 | 1.61 | -2.39 | 5.05 |
| All | 5.17 | 4.65 | 29.07 | -3.58 | 2.61 | -15.11 | 22.81 |

Table 6: Shares of income sources in 1996

| Sub-Group | Capital Income | Entrepr. Income | Wage Income | Social Assistance | Social Benefits | Direct Taxes | Disposable Income |
|---|---|---|---|---|---|---|---|
| Single adult | 0.78 | 0.30 | 7.66 | 2.01 | 1.25 | -3.27 | 8.74 |
| Pair, no child | 1.74 | 1.78 | 16.89 | 3.40 | 3.65 | -8.11 | 19.35 |
| Single Parent | 0.26 | 0.06 | 2.56 | 1.33 | 0.28 | -1.12 | 3.39 |
| Pair, children | 4.26 | 5.65 | 44.75 | 10.21 | 1.31 | -19.41 | 46.77 |
| Old Age | 1.53 | 0.19 | 0.08 | 3.05 | 6.65 | -2.06 | 9.44 |
| Other | 2.00 | 1.74 | 6.41 | 2.49 | 4.07 | -4.40 | 12.31 |
| All | 10.57 | 9.72 | 78.36 | 22.50 | 17.21 | -38.37 | 100.00 |

Table 7: Concentration coefficients in 1996

| | Social Assistance Concentration | | Direct Taxes Concentration | |
|---|---|---|---|---|
| Sub-Group | coefficient | difference | coefficient | difference |
| Single adult | -56.29 | -45.60 | 13.81 | -23.74 |
| | 5.95 | 6.60 | 5.66 | 6.18 |
| Pair, no child | -1.50 | 9.18 | 54.61 | 17.07 |
| | 7.13 | 7.75 | 2.75 | 3.70 |
| Single Parent | -60.07 | -49.39 | 6.40 | -31.15 |
| | 5.15 | 5.86 | 14.69 | 14.97 |
| Pair, children | -10.68 | . | 37.55 | . |
| | 2.25 | . | 1.80 | . |
| Old Age | -25.61 | -14.92 | 23.21 | -14.34 |
| | 4.92 | 5.65 | 8.82 | 9.21 |
| Other | 11.08 | 21.76 | 54.34 | 16.79 |
| | 6.23 | 6.87 | 4.99 | 5.64 |

The differences are calculated from the group, Pair with children.
Standard deviations are shown on the even rows.

Table 8: Within-group decomposition of the Gini coefficient in 1996

| Sub-Group | Capital Income | Entrepr. Income | Wage Income | Social Assistance | Social Benefits | Direct Taxes | Disposable Income |
|---|---|---|---|---|---|---|---|
| Single adult | 0.36 | 0.07 | 3.69 | -0.54 | -0.10 | -1.42 | 2.06 |
| Pair, no child | 0.80 | 0.92 | 6.10 | -0.66 | 0.27 | -3.06 | 4.37 |
| Single Parent | 0.17 | -0.01 | 1.10 | -0.20 | 0.19 | -0.53 | 0.72 |
| Pair, children | 2.28 | 2.26 | 12.84 | -2.15 | 0.18 | -6.24 | 9.18 |
| Old Age | 0.57 | 0.08 | 0.05 | 0.09 | 2.10 | -1.05 | 1.84 |
| Other | 1.16 | 0.78 | 2.18 | -0.16 | 0.99 | -1.79 | 3.16 |
| All | 5.17 | 4.65 | 29.07 | -3.58 | 2.61 | -15.11 | 22.81 |

To simplify the comparisons the rows have been multiplied by the income shares of the groups

Table 9: Residual terms in the within-group decomposition of the Gini in 1996

| | Capital Income | Entrepr. Income | Wage Income | Social Assistance | Social Benefits | Direct Taxes | Disposable Income |
|---|---|---|---|---|---|---|---|
| Residual | -0.18 | 0.56 | 3.12 | 0.03 | -1.02 | -1.03 | 1.48 |

Table 10: Change in the Gini table from 1990 to 1996

| Sub-Group | Capital Income | Entrepr. Income | Wage Income | Social Assistance | Social Benefits | Direct Taxes | Disposable Income |
|---|---|---|---|---|---|---|---|
| Single adult | 0.01 | -0.08 | 0.92 | -0.73 | -0.22 | -0.10 | -0.16 |
| Pair, no child | 0.48 | 0.57 | 0.01 | -0.23 | 0.69 | -0.90 | 0.56 |
| Single Parent | 0.06 | -0.02 | 0.43 | -0.48 | 0.07 | -0.13 | -0.01 |
| Pair, children | 1.11 | 1.20 | 6.41 | -0.77 | 0.21 | -3.09 | 5.03 |
| Old Age | 0.28 | -0.03 | 0.01 | 0.58 | 0.85 | -0.44 | 1.30 |
| Other | 0.40 | -0.77 | -7.64 | 0.07 | 1.01 | 2.01 | -4.95 |
| All | 2.34 | 0.87 | 0.14 | -1.57 | 2.61 | -2.66 | 1.76 |

Table 11: Testing for differences in the Gini table from 1990 to 1996

| Sub-Group | Capital Income | Entrepr. Income | Wage Income | Social Assistance | Social Benefits | Direct Taxes | Disposable Income |
|---|---|---|---|---|---|---|---|
| Single adult | 0 | 0 | 1 | -4 | 0 | 0 | 0 |
| Pair, no child | 1 | 0 | 0 | 0 | 2 | -1 | 0 |
| Single Parent | 0 | 0 | 0 | -4 | 0 | 0 | 0 |
| Pair, children | 1 | 2 | 4 | -3 | 0 | -4 | 4 |
| Old Age | 0 | 0 | 0 | 4 | 1 | -1 | 2 |
| Other | 0 | -2 | -4 | 0 | 2 | 4 | -4 |
| All | 3 | 0 | 0 | -4 | 4 | -4 | 2 |

The sign gives the direction of the change, and 1, 2, 3, 4, denote significance (two-sided test) with sizes 0.1, 0.05, 0.01, and 0.001, respectively.

Table 12: Change in the Gini table from 1976 to 1990

| Sub-Group | Capital Income | Entrepr. Income | Wage Income | Social Assistance | Social Benefits | Direct Taxes | Disposable Income |
|---|---|---|---|---|---|---|---|
| Basic education | -0.13 | -0.12 | -3.79 | -0.09 | -1.51 | 1.65 | -4.10 |
| Lower level | 0.30 | 0.64 | -0.12 | -0.54 | -0.18 | -0.32 | -0.29 |
| Upper level | 0.23 | 0.29 | 1.25 | -0.06 | -0.33 | -0.56 | 0.81 |
| Tertiary educ. | 0.20 | 0.11 | 0.11 | 0.06 | 0.13 | -0.10 | 0.53 |
| Higher degree | 0.41 | 0.05 | 3.00 | 0.34 | 0.44 | -1.07 | 3.26 |
| All | 1.00 | 0.97 | 0.45 | -0.30 | -1.44 | -0.40 | 0.20 |

Table 13: Changes in income shares from 1976 to 1990

| Sub-Group | Capital Income | Entrepr. Income | Wage Income | Social Assistance | Social Benefits | Direct Taxes | Disposable Income |
|---|---|---|---|---|---|---|---|
| Basic education | -1.15 | -6.48 | -22.44 | -1.34 | 2.15 | 6.86 | -22.20 |
| Lower level | -0.05 | 0.82 | 7.25 | 2.20 | 0.82 | -2.15 | 9.01 |
| Upper level | 0.02 | 0.60 | 4.88 | 1.25 | 0.42 | -1.48 | 5.87 |
| Tertiary educ. | 0.23 | 0.23 | 0.86 | 0.49 | 0.62 | -0.47 | 2.00 |
| Higher degree | 0.41 | 0.14 | 4.87 | 0.73 | 0.65 | -1.60 | 5.33 |
| All | -0.54 | -4.69 | -4.58 | 3.33 | 4.66 | 1.15 | 0.00 |

Table 14: Changes in population shares from 1976 to 1996

| Education | | | Family type | |
|---|---|---|---|---|
| Sub-Group | 1976-1990 | 1990-1996 | Sub-Group | 1990-1996 |
| Basic education | -22.19 | -5.96 | Single adult | 1.91 |
| Lower level | 10.21 | 0.75 | Pair, no child | 2.23 |
| Upper level | 6.30 | 1.76 | Single Parent | 1.46 |
| Tertiary educ. | 1.83 | 2.60 | Pair, children | 8.11 |
| Higher degree | 3.84 | 0.85 | Old Age | 0.76 |
| | | | Other | -14.46 |
| All | 0.00 | 0.00 | All | 0.00 |

Table 15: Change in the Gini table from 1990 to 1996

| Sub-Group | Capital Income | Entrepr. Income | Wage Income | Social Assistance | Social Benefits | Direct Taxes | Disposable Income |
|---|---|---|---|---|---|---|---|
| Basic education | 0.29 | -0.20 | -0.48 | 0.01 | 0.70 | -0.03 | 0.36 |
| Lower level | 0.42 | -0.01 | 1.86 | -0.93 | -0.02 | -0.63 | 0.72 |
| Upper level | 0.32 | 0.04 | -0.14 | -0.95 | 0.35 | -0.55 | -0.91 |
| Tertiary educ. | 0.05 | 0.37 | 0.37 | -0.12 | 0.22 | -0.44 | 0.42 |
| Higher degree | 1.26 | 0.68 | -1.47 | 0.41 | 1.36 | -1.01 | 1.18 |
| All | 2.34 | 0.87 | 0.14 | -1.57 | 2.61 | -2.66 | 1.76 |

Table 16: Change in the income shares from 1990 to 1996

| Sub-Group | Capital Income | Entrepr. Income | Wage Income | Social Assistance | Social Benefits | Direct Taxes | Disposable Income |
|---|---|---|---|---|---|---|---|
| Basic education | 0.47 | -1.81 | -7.23 | 1.64 | 1.68 | 0.40 | -5.14 |
| Lower level | 1.30 | -0.81 | -1.32 | 2.56 | 0.75 | -1.74 | 0.50 |
| Upper level | 0.76 | 0.05 | -0.95 | 1.83 | 0.66 | -1.67 | 0.47 |
| Tertiary educ. | 0.31 | 0.35 | 1.80 | 1.14 | 0.26 | -1.28 | 2.52 |
| Higher degree | 1.51 | 0.79 | -1.40 | 0.79 | 1.51 | -1.45 | 1.65 |
| All | 4.34 | -1.42 | -9.09 | 7.96 | 4.86 | -5.74 | 0.00 |