Discussion Papers No. 498, March 2007
Statistics Norway, Research Department

Magne Mogstad

Measuring Income Inequality under Restricted Interpersonal Comparability

Abstract:
The standard approach in empirical analyses of income distributions is to estimate income inequality in a country under the assumption of full interpersonal comparability of income. To be meaningful, this method requires that prices and qualities of goods as well as consumption habits are uniform across individuals in different regions of the country. In this paper, we pursue two alternative approaches to measure inequality under restricted interpersonal comparability of income. First, we estimate regional price indices, transform observed incomes into real incomes in an attempt to incorporate relevant non-income heterogeneity, and then aggregate across individuals. Second, we use the observed income data and account for non-income heterogeneity at the aggregation stage. Results based on a Norwegian register household panel data set indicates that both levels and trends in overall inequality as well as the inequality contributions of various income factors are robust to whether the income inequality analysis is based on the standard approach or the methods developed to cope with comparability problems within a country.

Keywords: Income inequality, restricted interpersonal comparability, heterogeneity, real income comparison, regional price index, decomposition of inequality by income factors

JEL classification: D31, D33, D63, I30

Acknowledgement: I am grateful to Rolf Aaberge for helpful discussions and suggestions, Tom Wennemo and Anders Nærø for computational assistance, and Ådne Cappelen, Erling Holmøy, Taryn Galloway, Audun Langørgen, Terje Skjerpen, and Steinar Strøm, and participants at the 29th General Conference of the International Association for Research in Income and Wealth for comments. The Norwegian Research Council has provided financial support for this project.

Address: Magne Mogstad, Statistics Norway, Research Department.
E-mail: magne.mogstad@ssb.no

Discussion Papers	comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

Abstracts with downloadable Discussion Papers
in PDF are available on the Internet:
http://www.ssb.no
http://ideas.repec.org/s/ssb/dispap.html

For printed Discussion Papers contact:

Statistics Norway
Sales- and subscription service
NO-2225 Kongsvinger

Telephone: +47 62 88 55 00
Telefax: +47 62 88 55 95
E-mail: Salg-abonnement@ssb.no

1. Introduction

An underlying assumption when making normative judgement on the basis of a set of income distributions according to the degree of inequality is that the assessment carries over to the distributions of economic welfare. This requires that there must be insignificant interpersonal variations in the conversion of individual income into individual welfare. Otherwise, an equal distribution of income may yield unequal welfare levels, and it becomes hard to defend equality in the income space as an expression of distributional justice. The reason is that income is a good that does not have an intrinsic value but is important merely as an instrument for individuals to pursue welfare. This implies that the population in a study of income inequality should, in principle, consist of income-recipients who are identical in every relevant aspect other than income. For this reason, comparisons of incomes across countries seek to adjust country currencies to common measures by accounting for important non-income differences such as variation in prices across countries; this is typically done by employing purchasing power parities.

Acknowledging, however, that the welfare basis of such real income comparisons may be rather limited due to methodological and data issues ranging from the basic index-number problem to disparities in national household surveys, cross-country studies of inequality regularly confine the comparisons to intra-country relative measures and do not seek absolute comparisons of levels of incomes in different countries.[1] By contrast, empirical analyses of income distributions within a country usually do not consider the implications of non-income differences between individuals beyond accounting for resource sharing and household economies of scale. Since empirical evidence suggests other important sources to comparability problems of incomes within a country, such as a substantial price difference of housing between urban and rural areas, intra-country relative measures may nevertheless suffer from a weak welfare basis. Consequently, the conventional assessments of income distributions within a country might be biased. As pointed out by Brandolini (2006), this bias is likely to carry over to assessments of absolute differences between countries, where one attempts to correct for differences in the average cost-of-living between countries, but not for differences in the cost-of-living across areas within the same country.

According to Coulter et al. (1992) and Cowell (1995), there are two strategies available for coping with problems of comparability in order to achieve distributions of income that mirror the

[1] Whilst Smeeding et al. (1993) and Atkinson et al. (1995) refrain from making absolute comparison of incomes across countries and limit the cross-country study exclusively to intra-country relative measures, Yotopoulos (1989), Schultz (1998), Milanovic (2002), and Sala-i-Martin (2006) attempt to assess the world income distribution and Atkinson (1996) and Brandolini (2006) provide EU-wide estimates of inequality. See Atkinson and Brandolini (2001) for a discussion of the issue of the comparability of the data commonly used to estimate the world income distribution.

distributions of unobserved welfare; either one transforms the income measure by incorporating the relevant non-income heterogeneity and aggregates across individuals or one uses the observed income data and accounts for non-income heterogeneity at the aggregation stage. In this paper, we will pursue both strategies. To this end, we utilise a register household panel data set covering the entire resident population of Norway in the period 1993-2001; these household and demographic data are supplemented with detailed income data from the Tax Assessment Files. Access to this unique data set allows us to account for non-income heterogeneity at a sufficiently disaggregate level. A necessary condition for either of the above strategies is that the heterogeneous population in question can be partitioned by relevant non-income differences into a set of mutually exclusive homogeneous subgroups; the incomes can thus be considered comparable within, but not necessarily between, the subgroups. Since the primary focus of this paper is on how to deal with comparability problems in the spatial dimension, we partition the Norwegian population - after using an equivalence scale to account for heterogeneity in household size and composition - into 90 subgroups according to the economic region to which they belong. The main criteria used for defining these regions are local labour market conditions, trade and service patterns as well as commuting and internal migration patterns. The purpose of this classification is to account for barriers to arbitrage within a country caused by transport costs, imperfect competition and information, and transaction costs of relocating and commuting. Since the capacity of individuals to purchase goods that are not perfectly tradable depends on the level of resources available to the people around them through the geographic pattern of competition, we would expect prices on certain goods, such as housing, to increase with the general income level in a region. In fact, we find a strong positive correlation between housing prices and the general income level across the economic regions. This indicates that the consumption potential of a given amount of income differs systematically between economic regions, which in turn suggests restricted interpersonal comparability of incomes. The positive correlation between income levels and housing prices across the economic regions is compatible with the prediction of the Harrod-Balassa-Samuelson proposition that the price levels on nontradable goods, and thus the overall consumer price level, tend to rise with country per capita income.

To achieve full comparability of incomes within the country, it is necessary to transform the observed incomes into real incomes according to relevant non-income differences in the population. Because the essential purpose of the income transformation carried out in this paper is to permit comparisons of income between individuals in different regions, we must be particularly interested in non-income differences that are systematically correlated with the general income levels across the regions. Such non-income differences could bias the results of the whole exercise of measuring inequality within a country. Thus, accounting for variation in housing prices across the

subgroups is critical for obtaining reliable estimates of real incomes. In fact, adjusting for geographical differences in the cost of housing in distributional analysis was one of the main recommendations made by the National Academy of Sciences Panel on Poverty and Family Assistance in the United States (Citro and Michael, 1995).[2] To this end, we use information about relative prices per square meter on houses sold in the various economic regions as a proxy for the relative prices on housing. Due to lack of credible data, we have to rely on an assumption of no variation in prices across the regions for all other goods. By estimating a set of regional price indices, we are able to transform observed incomes into real incomes. On the basis of the distribution of real incomes within the country, we may straightforwardly aggregate across individuals to achieve a measure of the overall inequality in the population. Although housing related expenditure accounts for about a quarter of aggregate household consumption expenditure in Norway (excluding mortgage payments), price differentials on other goods will also be expected to matter when we compare the consumption potential of the incomes of different individuals. However, this will be an argument against the proposed real income transformation only if there are important price differentials which are systematically different from the pattern of housing prices across the regions. Altogether, it is necessary to keep in mind the serious drawback of the conventional method of measuring inequality where price differentials within a country are simply ignored.

Unfortunately, neither regional price indices nor detailed information about the geographic price pattern of key goods at a sufficiently disaggregate level is always available in the OECD-countries. The binding constraint for transforming the observed incomes into real incomes may therefore very well be the data. Furthermore, it can be argued that inequality measures based on the estimates of real income suffer from a lack of welfare basis for reasons ranging from substitution bias in the price indices to systematic differences across regions in the qualities of goods offered as well as the consumption habits of individuals. For these reasons, we also reconsider the standard approaches conventionally applied at the aggregation stage of measuring inequality. Specifically, we propose a method that enables us to measure income inequality when incomes can be considered comparable within regions of a country, but we are unable to achieve adequate comparability between the regions by transforming the observed incomes into real incomes. On the basis of the distribution of income within each of the regions we estimate a set of region-specific inequality measures. By aggregating the inequality measures across the regions according to their population shares, we obtain an estimate of overall inequality where the term capturing inequality between/across regions is excluded. The purpose of this approach is to ensure a welfare basis of the measurement of inequality in a

[2] See Short et al. (1998), Short (2001), and Jolliffe (2006) for studies of poverty in the U.S. using an inter-area housing price index based on data on gross rent for apartments to adjust for geographical differences in the cost of housing.

heterogeneous population by restricting the comparison of incomes to individuals who live in the same region and are thus likely to face similar price pattern and qualities of goods as well as share consumption habits. This approach can also be interpreted as providing a measure of the average relative deprivation in a country along the lines of Runciman (1966), who argued that individuals compare themselves with a certain reference subgroup rather than with the population of the country as a whole.

When relying exclusively on intra-regional relative measures and not seeking absolute comparisons of levels of incomes between the regions, one runs the risk of disregarding genuine inter-regional differences. Whether or not to let region-specific inequality measures form the basis of an analysis of income inequality should thus not be conceived as a matter of the presence of relevant non-income differences within a country, but the relative size of the signal to the bias and the noise of relying on inter-regional information; comparison is a matter of degree rather than absolute.

The paper is organized as follows. Section 2 discusses the impact of non-income heterogeneity in the population on the choice of method for measuring inequality. In Section 3 we examine the extent to which levels and trends in income inequality as well as the results from decomposition analysis by income factor depend on the choice of method for measurement. Concluding remarks are provided in Section 4.

2. The impact of population heterogeneity on the choice of method for measuring income inequality

"Heterogeneity of factors that influence individual advantage is a pervasive feature of actual social evaluation. While we can decide to close our eyes to this issue by simply *assuming* that there is something homogeneous called 'the income' in terms of which everyone can be compared (and that variations of needs, personal circumstances, prices, etc, can be, correspondingly, assumed away) this does not resolve the problem – only evades it" (Sen and Foster, 1997, p204).

Arguably, the primary problem of measuring income inequality is not that of constructing indices from fundamental properties they presumably ought to have in the conventional context of a homogeneous population; it is a problem of the meaning that we can give to the measures we choose to employ subject to reasonable informational assumptions.[3] Indeed, the relevance of axiomatic results depends entirely on the meaningfulness of their constituent properties. Below, we demonstrate that the welfare basis of the conventional method for measuring inequality within a country is based on stringent assumptions about the relationship between observed income and economic welfare across

[3] For axiomatic characterisation of the Theil index, the Generalised Entropy family, and the Gini-coefficient see Foster (1983), Shorrocks (1984), and Aaberge (2001b), respectively.

the population; this may be a limitation to the practical relevance of the standard method. Alternative methods employing weaker and less controversial informational assumptions are subsequently considered. In end effect, the objective is to contribute to bridging the wide gap between theoretical work on the measurement of inequality, which presupposes a homogeneous population, and empirical counterparts forced to deal with a population of heterogeneous individuals inhabiting a heterogeneous environment.[4]

Before considering how to measure income inequality under restricted interpersonal comparability it is necessary to introduce some notation. Consider a population of n individuals and define for each person $i = 1,2,\ldots, n$:

Y_i - the equivalent income of person i after adjusting for household size and composition

p_i - the vector of prices facing person i after adjusting for differences in the qualities of goods

z_i - the vector of individual characteristics of person i

Let $u_i = v(Y_i; p_i, z_i)$ be the deterministic indirect utility function, which is an increasing function of Y_i. This function provides a measure of the economic welfare of individual i for equivalent income Y_i conditioned on the prices of goods p_i he is faced with and individual characteristics z_i. Note that the assumption of a common functional form implies that individuals have a common preference map and that any heterogeneity in needs and circumstances must be possible to define on the vector of individual characteristics. Thus, the vectors of individual characteristics are assumed to account for all interpersonal variation in the conversion of individual income into individual welfare for a given set of prices. Inverting u_i yields $Y_i = g(u_i; p_i, z_i)$, which measures how much income individual i would need to achieve welfare level u_i at the prices p_i given his individual characteristics z_i. We partition the population into r mutually exclusive and collectively exhaustive subgroups in which each member has identical z and faces the same p. Let the proportion of the population that belongs to subgroup j be $a_j = \frac{n_j}{n}$, where n_j represents the number of individuals in subgroup j, $j = 1, 2, \ldots, r$. Thus, $\sum_{j=1}^{r} a_j = 1$. If the entire population has the same p and z then $r = 1$. By contrast, if each individual differs in terms of z and/or p, then $r = n$.

2.1. Measuring income inequality in a homogeneous population: The conventional approach

Suppose that Y is perfectly measurable and fully comparable between individuals in the population. Furthermore, assume that z and p are homogeneous across the population, i.e. $r = 1$. In this benchmark

[4] See Foster (1984) for a discussion of the divergence between theoretical and applied work in the measurement of inequality.

case, the distribution of Y is cardinally equivalent to the distribution of welfare.[5] The overall inequality in the population based on the cumulative distribution function of Y can mechanically be expressed as a function of inequality within and between/across subgroups

$$I = \sum_{j=1}^{r} w_j I_j + R \tag{2.1}$$

where I is a country-specific measure of overall inequality that satisfies symmetry, replication invariance, mean independence and also the Pigou-Dalton transfer condition, I_j is I applied on subgroup j, w_j is the weight attributed to subgroup j, and the term R captures inequality between/across subgroups.[6] If the underlying assumption of full interpersonal comparability of income is satisfied, I will provide a meaningful summary measure of overall inequality in a population. Moreover, (2.1) enables us to study the relationship between overall inequality and inequality within and between/across subgroups of the population according to, say, gender or region of residence. By and large, the theoretical as well as the empirical literature on the measurement of inequality is based on this approach (Jenkins and Lambert, 1993). However, unless the population of study consists of homogeneous individuals inhabiting a homogeneous environment, this approach runs the risk of producing estimates of inequality that lack a welfare basis.

2.2. Measuring income inequality in a heterogeneous population: Transforming the income measure or adjusting the aggregation procedure?

In order for differences in welfare to arise exclusively from differences in income, as assumed in the previous section, all individuals must face the same prices and have the same individual characteristics. Below, we will consider how to measure income inequality when there is relevant non-income heterogeneity in the population.

2.2.1. Transforming the income measure

Empirical evidence suggests that prices of basic goods, such as housing, differ significantly between urban and rural areas within the same country. Under the assumption of perfect tradability - implying not only no transport costs, perfect competition, and complete information but also that individuals

[5] Formally, two measures are cardinally equivalent if the value of one measure can be obtained from the other by multiplying a positive constant and adding or subtracting another constant.

[6] For a strict statistical decomposition, the between-group inequality depends only on group means and the within-group inequality depends only on group inequality measures (Das and Parikh, 1982). As opposed to the inequality measures that are additively decomposable, the so-called generalised entropy family of inequality measures, the Gini-coefficient does not admit strict statistical decomposition into within- and between-group components but does also require an across term. For more on the subgroup decomposition issue, see e.g. Rao (1969), Shorrocks (1980, 1984), Cowell (1980, 1988), Das and Parikh (1982), Anand (1983), Lambert and Aronson (1993), Sen and Foster (1997), and Foster and Shneyerov (1999).

face no transaction costs to relocate or commute - the 'law of one price' will be satisfied for all goods (see e.g. Marris, 1984). In this case, observed price differences across subgroups should simply reflect differences in qualities of goods, i.e. each individual faces the same *p*. If there are barriers to arbitrage on the other hand, price differences will cause the welfare-equivalent of a given amount of income to vary systematically between subgroups within a country. This implies that $Y_i = g(u_i; p_i, z_i)$ may differ from $Y_k = g(u_k; p_k, z_k)$ even if $u_i = u_k$ and $z_i = z_k$.

Suppose that the *r* subgroups of the population have the same *z* but different *p*. To achieve full comparability of incomes within the country, a transformation of the observed incomes into real incomes is required. For a given subgroup *j*, the relationship between real income Y_j^* and observed income Y_j can be specified as

$$Y_j^* = Y_j \cdot c_{j,k}, \tag{2.2}$$

where $c_{j,k}$ is the cost-of-living index of an individual from a given reference-subgroup *k* relative to an individual from subgroup *j*. Let $c_{j,k}$ be defined as

$$c_{j,k} = \frac{e(u_k; p_k, z)}{e(u_k; p_j, z)}, \tag{2.3}$$

where the function $e(\cdot)$ gives the minimum expenditure level necessary to achieve a given utility level conditional on prices and individual characteristics. Thus, $c_{j,k}$ is given by the ratio of the minimum expenditure levels necessary to achieve the welfare level of reference-subgroup *k* at the prices facing subgroup *k* and *j*, respectively. Under the assumption of homothetic preferences, it follows that $e(u_k; p, z) = k(p)h(u_k; z)$.[7] Thus, $c_{j,k}$ is a function of observable prices only. As is well known, the cost-of-living index of subgroup *j* relative to subgroup *k* will then be bounded from below by the Paasche index denoted $P_{j,k}^P$ and from above by the Laspeyres index denoted $P_{j,k}^L$. These are defined as

$$P_{j,k}^L = \frac{\sum_{m=1}^{M} p_{m,k} q_{m,j}}{\sum_{m=1}^{M} p_{m,j} q_{m,j}} = \sum_{m=1}^{M} \left(p_{m,k} / p_{m,j} \right) s_{m,j}, \tag{2.4}$$

and

[7] A function is homothetic if it is an increasing transformation of a linearly homogeneous function. Intuitively, homotheticity implies that the indifference curves are radial copies of each other (see e.g. Sen and Foster, 1997).

(2.5) $$P_{j,k}^{P} = \frac{\sum_{m=1}^{M} p_{m,k} q_{m,k}}{\sum_{m=1}^{M} p_{m,j} q_{m,k}} = \left[\sum_{m=1}^{M} \left(p_{m,k} / p_{m,j} \right)^{-1} s_{m,k} \right]^{-1},$$

where the price and quantity of commodity $m = 1,2,\ldots,M$ supplied in subgroup j are denoted $p_{m,j}$ and $q_{m,j}$ respectively, whilst the expenditure share on commodity m in subgroup j is given by $s_{m,j} = p_{m,j} q_{m,j} \Big/ \sum_{m=1}^{M} p_{m,j} q_{m,j}$. $P_{j,k}^{L}$ can thus be written as an arithmetic average of the m price ratios $p_{m,k}/p_{m,j}$ that are weighted by subgroup j's expenditure shares, as seen from (2.4). In comparison, (2.5) show that $P_{j,k}^{P}$ can be written as a harmonic average of the same m price ratios that are weighted by subgroup k's expenditure shares.

Over the last decade, a consensus has emerged that bilateral comparisons should be made using a so-called superlative index number, one of which is the Fischer index defined as the geometric mean of $P_{j,k}^{L}$ and $P_{j,k}^{K}$ (Hill, 2004). In fact, for bilateral comparisons the superlative index numbers can be shown to be favourable from an axiomatic perspective and they are exactly equal to the true cost of living function for flexible expenditure functions when preferences are homothetic (Diewert, 1976, 1999). A problem with the bilateral indices (including the superlative ones) applied to a multilateral context, i.e. when $r > 2$, is that they are not in general transitive if the expenditure pattern varies across subgroups, as one would expect when prices vary across subgroups. Furthermore, the bilateral indices described above require information about the expenditure shares of at least one subgroup. In practice, however, we usually do not have information about subgroup-specific expenditure shares at a sufficiently disaggregate level. By contrast, the Symmetric Star method for multilateral comparisons compares subgroups indirectly via the average subgroup and thereby ensures transitivity.[8] Furthermore, if we apply the Average Basket version of the Symmetric Star method we only need information about the price ratios between the subgroups and the expenditure shares of the average subgroup, which corresponds to the expenditure shares used in the national consumer price index. According to the Average Basket method, the price index between subgroup j and k can be defined as

(2.6) $$\frac{P_k}{P_j} = \frac{P^{L}{}_{x,k}}{P^{L}{}_{x,j}} = \frac{\sum_{m=1}^{M} p_{m,k} q_{m,x}}{\sum_{m=1}^{M} p_{m,j} q_{m,x}} = \sum_{m=1}^{M} \left(p_{m,k} / p_{m,j} \right) s_{m,x}$$

[8] See Hill (1997) for a survey of multilateral methods for making comparisons of prices and quantities.

where $q_{m,x}$ is the quantity of commodity m consumed in the average subgroup x and $s_{m,x}$ the corresponding expenditure share. From (2.6) it is clear that the price index between subgroup j and k is defined as the ratio of the Laspeyres index for the average subgroup x and subgroup k and j, respectively.

Alternatively, we could have used the Average Price version of the Symmetric Star method and defined the price index between subgroup k and j as the ratio of the Paasche index for the average subgroup and subgroup k and j, respectively. To apply the Average Price method, however, information about subgroup-specific expenditure shares is required, although it would in this case suffice with commodity prices for the average subgroups rather than subgroup-specific commodity prices. In comparison, the Fischer Star index, which is a geometric mean of the pair of indices stemming from the Average Basket and Average Price method, requires information about subgroup-specific and average subgroup expenditure shares as well as subgroup-specific and average subgroup commodity prices.

Note that although the Average Basket method satisfies important index number tests, including the Weak Factor Reversal Test and the Average Test for price indices, it is subject to the adverse Gerschenkron effect (Hill, 1997). The Gerschenkron effect arises because expenditure patterns are likely to change in response to changes in relative prices, since individuals presumably substitute consumption towards relatively cheaper goods. As a result, measures of inequality across the subgroups based on real income measures constructed utilising (2.6) are likely to be upward biased if the expenditure pattern of the average subgroup is more similar to the expenditure pattern of the high-income subgroups compared to the low-income subgroups. Whilst the Average Price method is also subject to the Gerschenkron effect the Fischer Star method does not suffer from such bias, though it does require more information.

By replacing $c_{j,k}$ in (2.2) with P_k/P_j from (2.6) we can obtain estimates of real incomes Y^* for the entire population. Applied to the distribution of Y^*, the country-specific inequality index defined by (2.1) provides a measure of the overall inequality in real incomes as well as measures of inequality within and between/across subgroups. As long as the r subgroups of the population differ exclusively in p and the chosen price index is a reasonable approximation of the true cost of living index, the distribution of Y^* will correspond to the underlying distribution of welfare.

In addition to the Gerschenkron effect there are, however, several problematic aspects related to the welfare basis of the constructed real income measures. Firstly, the transformation of observed incomes to real incomes assumes that all commodities are available and identical in quality countrywide. For certain commodities, there are methods developed to account for quality differences, such as hedonic indices developed to adjust for quality differences in housing due to, say, location.

However, for a major class of commodities the quality and sometimes also the availability are inherently difficult to assess. This includes health care, education, and local government services, which are usually not provided by the market mechanism. If the quality and/or availability of such goods vary systematically within a country, then the real income measures will be biased. Furthermore, if preferences are non-homothetic there exists no unique, true measure of real income, since the cost-of-living index depends on the utility level of the reference-subgroup (Neary, 2004). Perhaps more importantly, the assumption of identical individual characteristics across the population is a crude one. Indeed, one could question whether the consumption habits of individuals apply broadly to the entire country or differ according to region of residence. Arguably, an individual's commodity requirements depend on the circumstances of his or her reference group, which are, in turn, presumably influenced by the community to which he or she belongs. If one agrees with Sen (1984) that there is significant variability in the commodity requirements within a given country, then the levels of welfare individuals can achieve for a given amount of income may vary depending on, say, their region of residence even when price patterns and qualities of goods across regions are the same.

2.2.2. Adjusting the aggregation procedure

Suppose that the r subgroups of the population differ in terms of z and/or p and that we are not able to achieve adequate comparability of incomes between the subgroups by a transformation of the observed incomes into real incomes. On the basis of the r subgroup-specific estimates of the inequality measures in (2.1) one may obtain a complete ordering over a set of possible distributions of income for each subgroup. The reason is that the population of each subgroup consists of identical individuals in every relevant aspect other than income. Since the incomes are not comparable between the subgroups, the subgroup aggregation in (2.1) is, as pointed out in Coulter et al. (1992) and Cowell (1995), contentious. The problem is twofold:

- The measure of between/across-group inequality is based on non-comparable elements.
- The weights of the within-group inequalities, which conventionally depend on the income shares of each subgroup, are no longer appropriate.

Although measured inequality within each subgroup can be seen as reflecting a genuine disparity among individuals' abilities to achieve welfare, the between/across-group measure of inequality does not necessarily capture differences in the welfare of individuals belonging to different subgroups. The reason is that prices and individual characteristics vary between subgroups implying that between/across-group inequality looses its information value for assessing overall inequality; we are

comparing apples with oranges. Accordingly, the weighting scheme of within-group inequalities cannot be based on the income shares of the subgroups. Thus, it is necessary to introduce an alternative weighting scheme that does not involve subgroup-specific average incomes when forming an overall measure of inequality based on the subgroup-specific inequality estimates. We propose setting the weights according to the population shares of the subgroups. By inserting a_j for w_j and dropping R in (2.1), the inequality in the population as a whole $\tilde{I}$ can be expressed exclusively as a weighted sum of inequality within the subgroups

$$\tilde{I} = \sum_{j=1}^{r} a_j I_j \ . \tag{2.7}$$

Above, $\tilde{I}$ is justified as a measure of overall income inequality in a heterogeneous population when incomes can be considered comparable within subgroups of a country, but we are not able to achieve adequate comparability between the subgroups by transforming the observed incomes into real incomes.[9] Alternatively, $\tilde{I}$ can be interpreted along the lines of the strand of literature on relative deprivation starting with Runciman (1966), who argued that individuals may compare themselves with some reference group within the society rather than with the society as a whole. From Kakwani (1984) it follows that when I_j is represented by the Gini-coefficient, $\tilde{I}$ can be viewed as a measure of the average relative deprivation suffered by the population when individuals compare their incomes within their own subgroups only and therefore do not feel deprived relative to members of other subgroups.[10]

Note that in terms of informational requirements, the frequently used dominance criteria to rank income distributions in a heterogeneous population may be considered as an intermediate between the income transformation approach and the method based on subgroup-specific inequality measures. For example, unlike the income transformation approach, application of the sequential dominance conditions suggested by Atkinson and Bourguignon (1987) do not involve cardinal specification of transformation scales.[11] On the other hand, the dominance test requires not only partitioning of a heterogeneous population into subgroups distinctive in terms relevant non-income differences, but also that the subgroups can be unambiguously ranked according to the welfare-

[9] On the basis of a similar argument, Mogstad et al. (2007) proposes to measure poverty within a country on the basis of a set of subgroup-specific poverty lines rather than on the basis of a joint country-specific poverty line.

[10] Berrebi and Silber (1985) show that the income inequality measures commonly used can all be expressed as an income weighted sum of individual deprivation coefficients, distinctive in terms of the way deprivation is defined. See e.g. Yitzhaki (1979) for an alternative measure of relative deprivation expressed as the product of the mean income and the Gini-coefficient for a society or a subgroup thereof.

[11] This method has been extended by Atkinson (1992), Jenkins and Lambert (1993), Chambaz and Maurin (1998), and Lambert and Ramos (2002) to deal with changing demographics, poverty, and the principle of diminishing transfers. See e.g. Fleurbaey et al. (2003) for dominance conditions concerned with the robustness of the assessments of income distributions to the choice of cardinalisation of transformation scales.

equivalents of a given amount of income. This is not required by the method based on the set of subgroup-specific inequality measures. While it may be possible to agree on a suitable ranking of subgroups by a single source of heterogeneity, say, the relative needs of household types or price differentials between regions, a widespread agreement is less likely to be established about the ranking of subgroups by multiple sources of heterogeneity. In fact, if subgroups differ along several non-income dimensions, ranking of the subgroups requires information about the relative impact of these differences on individuals' welfare levels. Specifically, one is forced to provide answers to questions such as 'For a given amount of income, is a couple without children living in a region with relatively high prices needier than a couple with children facing lower prices?'. Unless consensus about ranking is established in such cases, dominance tests cannot solve the problem of assessing income distributions in a heterogeneous population. Instead, we have to rely on either the income transformation approach or the method based on a set of subgroup-specific inequality measures.

3. The sensitivity of income inequality estimates to the assumptions of interpersonal comparability

The objective of this section is to assess the sensitivity of estimates of income inequality to the choice between the standard method, which presupposes a homogeneous population, and the methods developed to cope with comparability problems within a country. Specifically, we apply the methods discussed in Section 2 to examine the extent to which levels and trends in income inequality depend on whether the basis of analysis is (i) a country-specific inequality measure based on observed incomes, (ii) a country-specific inequality measure based on real incomes or (iii) a set of region-specific inequality measures. A major discussion in the literature on income distribution revolves around the manner in which various income factors contribute to inequality in disposable income. For example, the policymaker may be interested in assessing the extent to which inequality is due to earnings or capital income or studying the redistributive nature of taxes and transfers. We will therefore also examine the sensitivity of results from a decomposition analysis by income factors to the choice of method. In fact, the results from the decomposition analysis are of interest in their own right as no decomposition analysis has so far been carried out for Norway in the period of time we are considering.

3.1. Data and methodological assumptions

The empirical analysis is based on a register household panel data set covering the entire resident population of Norway in the period 1993-2001. Access to this data set allows us to account for non-income heterogeneity at a sufficiently disaggregate level. By contrast, the data sets used in most cross-

country comparisons of income inequality, such as the Luxembourg Income Study database, do not contain enough observations to deal with heterogeneity within a country in a sound manner. Indeed, Aaberge (2001a) demonstrates that when sampling errors are taken into account, the complete ranking of countries suggested in the OECD study by Atkinson et al. (1995) have to be replaced by a ranking of countries in a few groups. The register panel data set with household and demographic information is supplemented with detailed income data from the Tax Assessment Files, which are collected from tax records and other administrative registers rather then interviews and self-reporting methods. The coverage and reliability of Norwegian income registers are considered to be very high, as is documented by the fact that the quality of such national datasets of income received the highest rating in a data quality survey in the Luxembourg Income Study database (Atkinson et al., 1995).

In this paper, we use disposable income as the focal variable for the empirical analysis of income inequality. Disposable income, which is defined in close agreement with international recommendations (e.g. Expert Group on Household Income Statistics, 2001), incorporates earnings, self-employment income, capital income, transfers and taxes. To enable the comparison of incomes across individuals belonging to households of varying size and composition the OECD equivalence scale is applied; the weight of the first adult in the household is set to 1, each additional adult is given a weight of 0.7, and each child gets a weight equal to 0.5. The robustness of the inequality estimates to the choice of equivalence scale is examined by the use of the EU equivalence scale, which gives the first adult the weight 1, each additional adult is given the weight 0.5, and each child the weight 0.3.

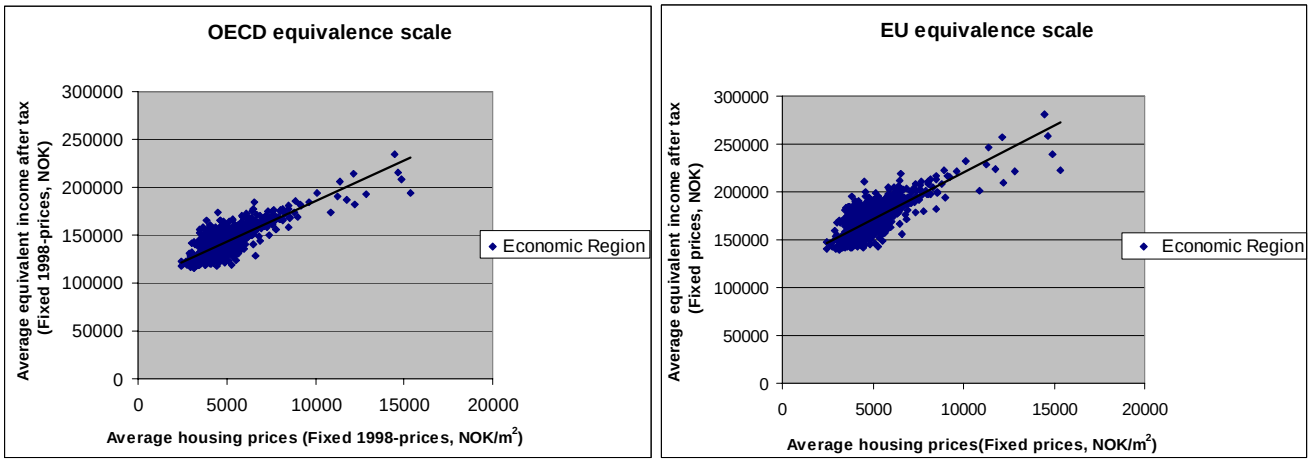
As previously addressed, a necessary condition for the income transformation approach as well as the method based on subgroup-specific inequality measures is that the population can be partitioned by relevant non-income differences into a set of mutually exclusive and collectively exhaustive subgroups that are homogeneous. In this case, incomes can be considered to be comparable within but not necessarily between the subgroups. Because the focus of the empirical analysis is to deal with comparability problems in the spatial dimension, we partition the population into 90 subgroups determined by the economic region of residence. The main criteria used for defining the economic regions are local labour market conditions, trade and service patterns as well as commuting and internal migration patterns.[12] The main goal of this classification is to account for barriers to arbitrage within a country caused by transport costs, imperfect competition and information, and transaction costs of relocating and commuting. Specifically, these economic regions constitute a regional level between the 19 counties and the 435 municipalities in Norway. Since the capacity of

[12] To supplement these formal criteria forming the basis for the definition of economic regions, geographical expertise accumulated in Statistics Norway as well as local knowledge from the municipalities have been utilised. See Statistics Norway (2000) for a detailed description of the classification of economic regions, which roughly correspond to the NUTS 4 – level in EU's regional classification.

individuals to purchase goods that are not perfectly tradable depends on the level of resources of the other people around them due to the geographic pattern of competition, we would expect prices on certain goods, such as housing, to increase with the general income level in a region. In fact, Figure 1 shows a correlation of 0.79 between average housing prices and the average equivalent income across the economic regions, independent of the choice of equivalence scale. This indicates that the consumption potential of a given amount of income differs systematically between economic regions, which in turn suggests restricted comparability of incomes within the country.

The positive correlation between income levels and housing prices across the economic regions is compatible with the prediction of the Harrod-Balassa-Samuelson proposition that price levels on nontradable goods tend to rise with country per capita income.[13] The basic argument underlying this proposition is that rich countries appear to be relatively more productive in tradables than nontradables. If the law of one price holds in the tradable sector, then cross-country relative wages are determined by productivity differences in tradables. In rich countries, the producers of nontradeable goods must set their prices relatively high to match the high wages in the tradeable sector. This implies that both nontradables as well as a representative basket of nontradable and tradable goods will be more expensive in high-income countries compared to low-income countries.

Figure 1: Average housing prices and average income level by economic region in Norway, 1993-2001 (Fixed 1998-prices, NOK)



Because the essential purpose of the income transformation carried out in this paper is to permit comparison of incomes between individuals in different regions, we must be particularly interested in non-income differences that are systematically correlated with general income levels across the

[13] See e.g. Rogoff (1996) for an introduction to the Harrod-Balassa-Samuelson theory and a survey of cross-country studies providing substantial support for the Harrod-Balassa-Samuelson proposition.

regions. Such non-income differences could bias the result of the whole exercise of measuring inequality within a country. Thus, accounting for variation in housing prices across the regions is critical for obtaining reliable estimates of real incomes. To this end, we use information about relative prices per square meter on detached houses sold in the various economic regions as a proxy for the price ratios for housing related goods in the price index defined by (2.6).[14] For all other goods we assume no variation in prices across the regions, since we lack credible data sources. Furthermore, a common presumption is that housing prices are positively correlated with prices on non-housing goods.[15] If so, the estimated differences between real incomes and observed incomes will be a lower bound estimate of the true differences between real incomes and observed incomes. What remains in order to use (2.6) to achieve estimates of real incomes is to determine the weights for housing versus non-housing goods for the average region or equivalently for the country as a whole. To this end, we use data on expenditure shares obtained from the Norwegian household budget survey; these data also form the basis for determining the weights in the national consumer price index.[16] In the period 1993-2001, the shares of housing related expenditures in aggregate household consumption expenditure range from about 22 to 26 per cent (excluding mortgage payments).[17] Since there are 90 regions and we have access to annual information about prices and expenditure shares on housing for 9 years, altogether 810 regional price indices are estimated on the basis of (2.6). The maximum discrepancy in the price indices turns out to be 0.24.

Note that this paper does not assume that housing prices are the same within each region, only that there are no barriers to arbitrage within the regions. Indeed, there are considerable price differences on housing also within certain regions, in particular among the various districts of the capital-region Oslo. However, as long as individuals are not required to live in certain areas of a region due to family situation or to participate in the local labour market, individuals from the same region will face the same prices although the prices on the goods they actually consume may vary due to, say, differences in purchasing power.

---

[14] One could argue that it would be more appropriate to use rental prices rather than real estate prices. However, detailed data on local level for rental prices are not available in Norway. Moreover, most people in Norway are, by large, owners rather then renters.

[15] For example, this assumption is made in Short et al. (1998), Short (2001) and Jolliffe (2006) when studying the sensitivity in the spatial distribution of poverty in the U.S. to cost of living adjustments.

[16] The household budget survey is based on personal interviews and detailed accounting in a representative sample of private households across the country. See Statistics Norway (2002) for a detailed description of the household budget survey.

[17] Source: Division for Economic Indicators, Statistics Norway

3.2. Definition of Inequality Measures

To summarize the informational content of the Lorenz curve and to achieve rankings of intersecting Lorenz curves, the conventional approach is to employ the Gini-coefficient. To examine the extent to which the empirical results depend on the choice of inequality measure, the conventional approach is to complement the Gini-coefficient with measures from the Atkinson or Theil family. However, the Gini-coefficient and inequality measures from the Atkinson or the Theil family have distinct theoretical foundations which make it inherently difficult to evaluate their capacities as complimentary measures of inequality. As demonstrated by Aaberge (2000, 2007), an alternative approach for examining inequality in the distribution of income is to rely on the moments of the scaled conditional mean curve *M* defined by

$$M(u)=\frac{E\left[Y \mid Y \le F^{-1}(u)\right]}{\mu}=\begin{cases}\dfrac{1}{u\mu}\displaystyle\int_0^u F^{-1}(t)dt, & 0<u\le 1\\ 0, & u=0\end{cases} \tag{3.1}$$

where F is the observed country-specific cumulative distribution function with mean μ, and F^{-1} its left inverse.[18] Specifically, the first, the second, and the third order moments of the scaled conditional mean curve prove to make up a fairly good summarisation of the conditional mean curve as well as the Lorenz curve. The kth order moment of the scaled conditional mean curve for F, denoted $D_k(F)$, is defined by

$$D_k(F)=\int_0^1 ku^{k-1}\left(1-M(u)\right)du, \quad k=1,2,\ldots \quad . \tag{3.2}$$

Aaberge (2007) shows that whilst the second moment can be represented by the Gini-coefficient (*G*), the first moment can be represented by the Bonferroni coefficient (*B*). The Bonferroni coefficient emphasises on changes that take place in the lower part of the distribution. The third moment can be represented by an inequality measure (*C*) that focuses on changes that takes part in the upper part of the income distribution. In this paper, we will examine the sensitivity of the empirical results to the choice of inequality measure by complementing the information provided by *G* with its two close relatives *B* and *C*.[19] Hence, we meet the most common criticism of the Gini-coefficient, namely that is insensitive to redistribution of income at the lower end of the distribution (see. e.g. Wiles, 1974)

[18] For a given u, *M(u)* is the ratio of the mean income of the poorest 100u per cent of the population and the overall mean. By inserting for the Lorenz curve in (3.1) it follows straightforwardly that the scaled conditional mean curve is a representation of inequality that is equivalent to the Lorenz curve.

[19] Since *G*, *B*, and *C* have a common theoretical foundation and proves to complement each other with regard to sensitivity to transfers, Aaberge (2007) treats them as a group called Gini's Nuclear Family.

When full interpersonal comparability of observed incomes can be justified, the *G*, *B*, and *C* coefficients for overall inequality in the distribution *F* of observed income Y can be expressed as

(3.3) $$G = \frac{1}{\mu}\int_0^\infty F(y)\big(1-F(y)\big)dy\,,$$

(3.4) $$B = -\frac{1}{\mu}\int_0^\infty F(y)\log F(y)dy\,,\text{ and}$$

(3.5) $$C = \frac{1}{2\mu}\int_0^\infty F(y)\big(1-F^2(y)\big)dy\,.$$

By replacing the distribution *F* of observed income *Y* with the distribution *F** of real incomes *Y** (with mean μ^*), the corresponding measures of overall inequality can be defined by

(3.6) $$G^* = \frac{1}{\mu^*}\int_0^\infty F^*(y)\big(1-F^*(y)\big)dy\,,$$

(3.7) $$B^* = -\frac{1}{\mu^*}\int_0^\infty F^*(y)\log F^*(y)dy\,,\text{ and}$$

(3.8) $$C^* = \frac{1}{2\mu^*}\int_0^\infty F^*(y)\big(1-F^{2^*}(y)\big)dy\,.$$

In the case where incomes are considered to be comparable within but not between the regions, it follows from (2.7) that measures of overall inequality in the population can be defined as the weighted average of the region-specific inequality measures. By employing (3.3)-(3.5) to region *j's* cumulative distribution function F_j (with mean μ_j), the measures of overall inequality corresponding to (2.7) can be defined by

(3.9) $$\tilde{G} = \sum_{j=1}^{r} a_j G_j = \sum_{j=1}^{r}\frac{a_j}{\mu_j}\int_0^\infty F_j(y)\big(1-F_j(y)\big)dy\,,$$

(3.10) $$\tilde{B} = \sum_{j=1}^{r} a_j B_j = -\sum_{j=1}^{r}\frac{a_j}{\mu_j}\int_0^\infty F_j(y)\log F_j(y)dy\,,\text{ and}$$

(3.11) $$\tilde{C} = \sum_{j=1}^{r} a_j C_j = \sum_{j=1}^{r} \frac{a_j}{\mu_j} \int_0^{\infty} F_j(y)\left(1 - F_j^2(y)\right) dy ,$$

where a_j is region *j's* population share.

Note that these rank-dependent measures of inequality can be decomposed in a way that provides a direct link between inequality measures based on the standard method presupposing a homogeneous population and inequality measures based on the methods developed to cope with cases characterised by restricted interpersonal comparability of incomes. For example, the Gini-coefficient G defined by (3.3) admits the following decomposition

(3.12) $$G = \tilde{G} + R = G^* + \left(R - R^*\right),$$

where R and R^* are terms that capture inequality between/across regions in observed and real incomes, respectively, and G^* and $\tilde{G}$ are defined by (3.6) and (3.9).[20] The above decomposition is attractive since it provides a direct link between G, G^*, and $\tilde{G}$, thus it allows us to straightforwardly determine, say, the contribution of $\tilde{G}$ to G.[21] Moreover, the decomposition demonstrates that G will differ from $\tilde{G}$ when there is substantial inequality between/across regions in observed incomes and that G deviates from G^* insofar as there are significant differences in between/across inequalities in real and observed incomes.

3.3. Decomposition of Inequality Measures by Income Factors

Assume that there are l = 1,2,…, s mutually exclusive and collectively exhaustive sources of income, such that $Y = \sum_{l=1}^{s} Y_l$. The scaled conditional mean curve defined by (3.1) may then be expressed as

(3.13) $$M(u) = \sum_{l=1}^{s} \frac{\mu_l}{\mu} \frac{E\left[Y_l \mid Y < F^{-1}(u)\right]}{\mu_l} \quad ,$$

---

[20] Following Rao (1969), R can be expressed as $R = \sum_{j=1}^{r-1} \sum_{k=j+1}^{r} \left\{ a_j \frac{a_k \mu_k}{\mu} \left(R_{jk} - G_j\right) + a_k \frac{a_j \mu_j}{\mu} \left(R_{kj} - G_k\right) \right\}$, where $R_{lm} = \frac{1}{\mu_l} \int_0^{\infty} F_l(Y)\left(1 - F_m(Y)\right) dY$. The term R^* is obtained by a similar decomposition applied to F^*.

[21] Alternative subgroup decompositions of the Gini-coefficient are proposed in Bhattacharya and Mahalanobis (1967), Pyatt (1976) and Aaberge et al. (2005). More on the derivation and interpretation of the subgroup decomposition of the Gini-coefficient, see Das and Parikh (1982), Silber (1989), Yitzhaki and Lerman (1991), Lambert and Aronson (1993), Yitzhaki (1994) and Dagum (1997).

where the factor income share μ_l/μ is the ratio between the means of income factor l and disposable income. When full interpersonal comparability of observed incomes can be justified we may, following Rao (1969), express $D_k(F)$ as the sum of inequality contributions from the income factors to overall inequality in disposable income. Specifically, by inserting (3.13) into (3.2) yields

$$D_k(F) = \sum_{l=1}^{s} \beta_{lk}, k = 1, 2, \ldots \quad , \tag{3.14}$$

where the inequality contribution β_{lk} of income factor l is given by the product of the factor income share and the concentration coefficient γ_{lk} of factor l

$$\beta_{lk} = \frac{\mu_l}{\mu} \gamma_{lk}, \tag{3.15}$$

and

$$\gamma_{lk} = \int_0^1 k u^{k-1} \left(1 - \frac{E\left[Y_l \mid Y \le F^{-1}(u) \right]}{\mu_l} \right) du, \quad k = 1, 2, \ldots \quad . \tag{3.16}$$

The factor concentration coefficient γ_{lk} can be interpreted as the conditional inequality measure of factor l given the rank order in disposable income. Thus, a negative value of γ_{lk} implies that income factor l gives an equalizing contribution to overall inequality in disposable income. If each individual receives an equal amount of income factor l then $\gamma_{kl} = 0$. This demonstrates that $D_k(F)$, and thus G, B, and C defined by (3.3)–(3.5), can be expressed as weighted averages of factor concentration coefficients with factor income shares as weights.[22]

By employing (3.14) to the distribution of real disposable income it follows that $D_k(F^*)$, and thus G^*, B^*, and C^* defined by (3.6)–(3.8), can be expressed as the weighted average of factor concentration coefficients with factor real income share as weights

$$D_k\left(F^*\right) = \sum_{l=1}^{s} \beta_{lk}^*, k = 1, 2, \ldots \quad , \tag{3.17}$$

[22] Note that the decomposition method defined by (3.15) provides a simultaneous treatment of the income factors in question. This implies that we focus on the contributions from the various income factors to the observed overall income inequality. By contrast, the elasticity approach for decomposing the Gini-coefficient by income sources proposed by Lerman and Yitzhaki (1985) concerns the effect of a marginal change in an income component. See also Fei et al. (1978), Pyatt et al. (1980), Kakwani (1980), Silber (1989, 1993), and Yao (1997) for alternative decomposition methods of the Gini-coefficient.

where β_{lk}^{*} is the inequality contribution of real income factor *l* defined by (3.15) and (3.16) employed to real disposable income.

A similar factor decomposition of $\tilde{G}$, $\tilde{B}$ and $\tilde{C}$ defined by (3.9)-(3.11) is obtained by employing (3.14) for each region-specific inequality measure

$$\tilde{D}_k\left(F_1, F_2, ..., F_r\right) = \sum_{l=1}^{s} \tilde{\beta}_{lk}, k = 1, 2, ... \quad , \tag{3.18}$$

and

$$\tilde{\beta}_{lk} = \sum_{j=1}^{r} a_j \frac{\mu_{jl}}{\mu_{j+}} \gamma_{jlk} \tag{3.19}$$

where μ_{jl}/μ_{j+} is the ratio between the means of income source *l* and disposable income in region *j*, i.e. the income share of factor *l* in region *j*. The region-specific factor concentration coefficient γ_{jlk} can be considered a measure of the interaction between income factor *l* and disposable income in region *j*. By taking the weighted sum of the products of the region-specific income shares and concentration coefficients of factor *l*, we get the inequality contribution of factor *l* to overall inequality in disposable income, $\tilde{\beta}_{lk}$.

3.4. Empirical results

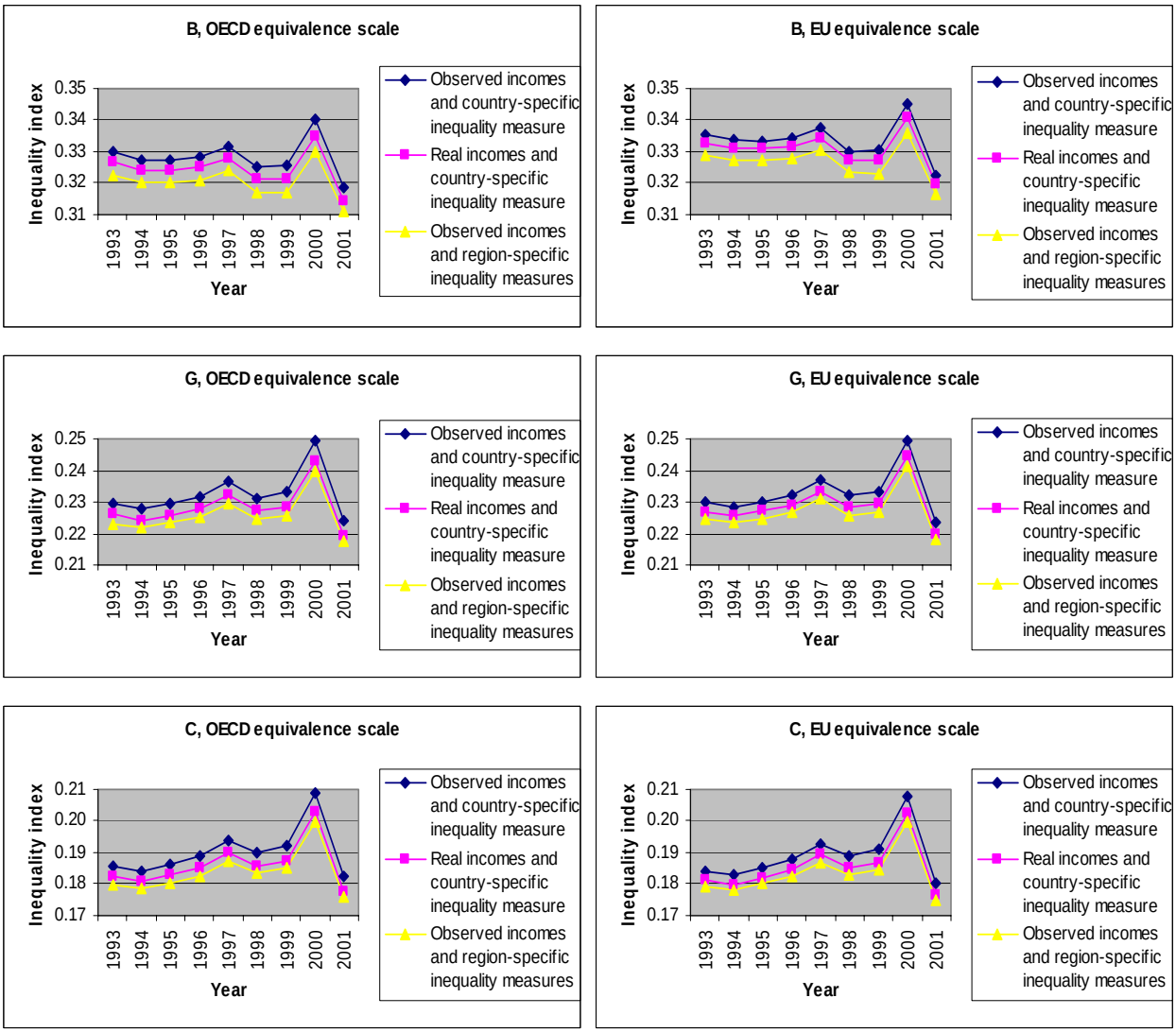
Figure 2 displays estimates of income inequality from the standard method based on observed incomes and country-specific inequality measures defined by (3.3)–(3.5). Moreover, this figure shows inequality estimates based on real incomes and country-specific inequality measures defined by (3.6)–(3.8), and inequality estimates based on observed incomes and region-specific inequality measures defined by (3.9)-(3.11). The results demonstrate that the levels and trends in income inequality are largely unaffected by the choice between the standard method and the methods developed to cope with comparability problems within a country. Specifically, estimates of income inequality based on country-specific inequality measures applied to observed incomes are only slightly larger than the inequality estimates based on country-specific inequality measures applied to real incomes. The reason for the small differences in inequality estimates is that the contribution of inequality within the regions to overall inequality, both cross-sectionally and to the trend over time, turns out to be substantially more important than the contribution of inequality between/across regions. Since estimates of inequality within regions are independent of whether we rely on observed or real incomes, estimates of overall inequality will only be slightly affected by the choice between these two approaches. Thus, we

may infer that although variation in housing prices across regions leads to considerable differences between observed incomes and real incomes, these differences do not carry over to estimates of overall inequality. Moreover, inequality estimates based on a country-specific inequality measure applied to observed incomes are only slightly larger than those based on a set of region-specific inequality measures where the between/across measures of inequality are excluded.

Note that the level of and trend in income inequality displayed by Figure 2 correspond, by and large, with results from other studies of income inequality in Norway in the 1990s based on survey data, such as Fjærli and Aaberge (2000). Unlike studies of relative poverty, no anti-cyclical relationship between the business cycle and the degree of inequality can be identified.[23] In fact, the degree of inequality is, in general, stable during the boom in the Norwegian economy, which lasted from 1993 to 1998. In this period, the economy was characterized by reduced unemployment rates, increased participation in the work force, and high growth in real wages. After the peak in 1998, the following years were characterized by a small downturn with somewhat weaker labour market and lower economic growth. As pointed out by Fjærli and Aaberge (2000), the rise in income inequality in 2000 should not be interpreted as a result of the business cycle, but largely driven by a tax reform that affected the financial incentives in the corporate sector and the income shifting incentives in small enterprises. The decomposition analysis by income factors carried out below draws further light on the reasons for the perhaps surprisingly stable income distribution in Norway in the period 1993-2001.

[23] See Galloway and Mogstad (2006) for an anti-cyclical relationship between the business cycle and the percentage poor in Norway from 1993-2001. In comparison, Aaberge et al. (2000) find a relatively stable income distribution over the recession during the late 1980s and early 1990s when unemployment rose dramatically in Norway.

Figure 2. Country-specific and region-specific inequality measures based on observed and real incomes, 1993-2001



Measures of inequality, such as the Gini-coefficient, do not, however, offer an immediate interpretation that clarifies the significance of a certain change in inequality. In general, neither the numerical values of inequality measures nor the numerical values of changes in inequality measures have any straightforward meaningful interpretation per se, and are primarily used as means to compare and order distributions by degree of inequality. Thus, a method for quantifying the economic implications of the observed differences in inequality is required to draw conclusions about the actual impact of relying on the methods developed to cope with comparability problems within a country rather than using the standard method presupposing a homogeneous population. To this end, we apply a method for interpreting changes in rank-dependent measures of inequality suggested by Aaberge

(1997). By means of a hypothetical intervention of a tax/transfer reform, this method offers an intuitive appealing interpretation of the magnitude of the differences in inequality estimates depending on the choice of method for measurement. Suppose, for example, that for a given year the estimated G is 10 per cent larger than the estimated G^*. This difference in inequality corresponds to introducing a proportional tax with tax rate equal to 10 per cent of individuals Y, and then redistributing the collected tax revenue as equal-sized amounts to the individuals. Such an intervention would, of course, leave the mean income unchanged. Figure 3 shows that the estimated differences in the inequality range from about 1 to 5 per cent, depending on the choice of method for measurement. Thus, we may infer that the choice of method for measuring inequality is economically of modest consequence on the estimated level of and trend in income inequality in Norway.

Figure 3. Percentage differences in country-specific and region-specific inequality measures based on observed and real incomes, 1993-2001

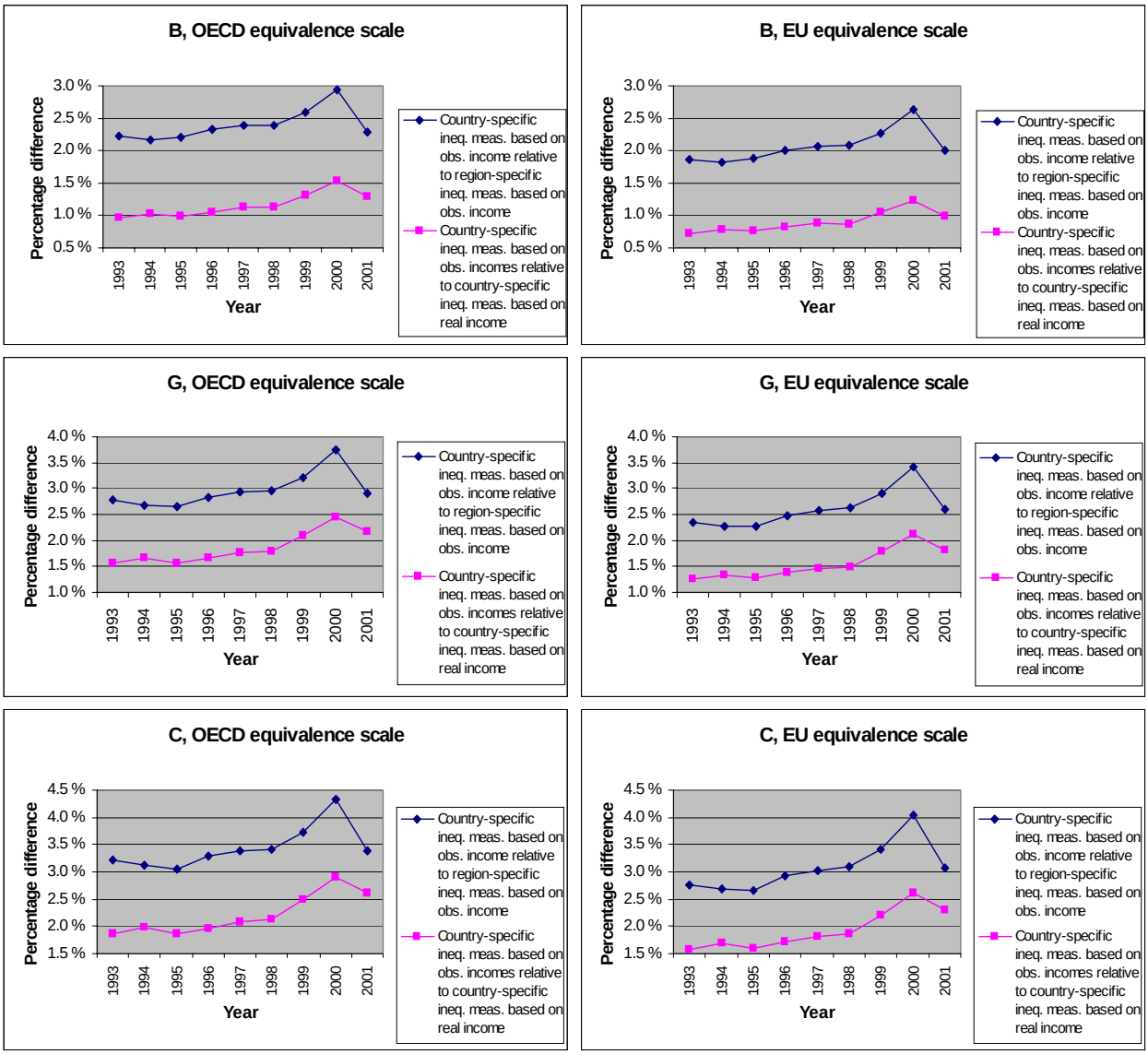
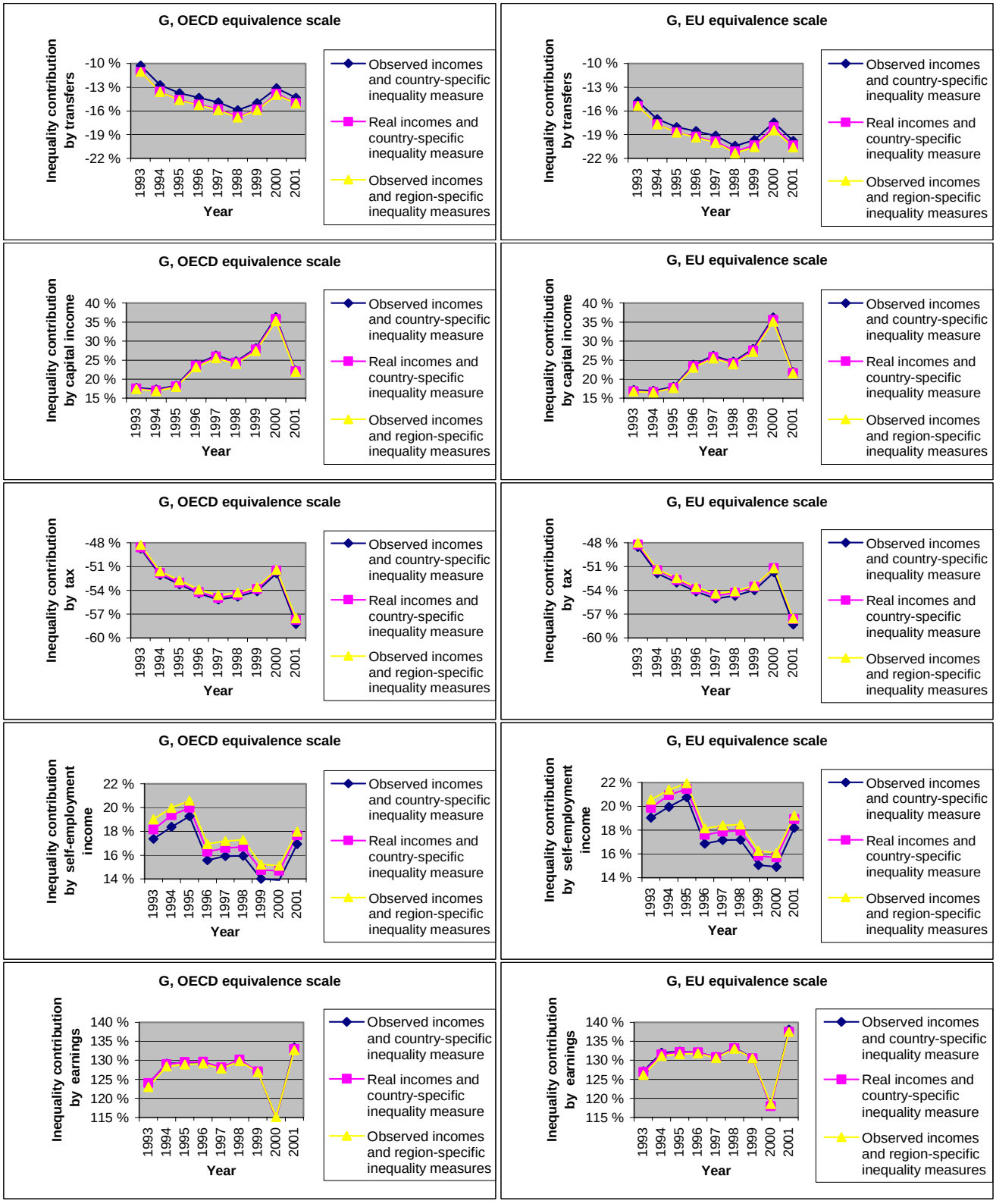


Figure 4. Contribution by income factor to inequality in disposable income measured by the Gini-coefficient when country-specific and region-specific inequality measures based on observed and real incomes form the basis, 1993-2001



Although the levels and trends in inequality in Norway appear to be robust with respect to the choice of method for measurement, it is not obvious how a decomposition analysis by income factors will be affected by this methodological choice. Figure 4 shows the contribution of various income factors to inequality in disposable income measured by the Gini coefficient, when the analysis is based on country-specific inequality measures applied to observed and real incomes as well as region-specific inequality measures based on observed income.[24] It is clear that the inequality contributions of various income factors also are rather robust to the choice of method for measurement. As expected, earnings are shown to be the dominating income factor, which clearly increases inequality in the distribution of disposable income. The strong disequalising effect of earnings is driven by its high share of disposable income. In fact, the concentration coefficients of earnings are relatively low compared to the concentration coefficients of income from capital and self-employment. In comparison, taxes, and to a less extent transfers, have equalising effects on the distribution of disposable income. The tax and transfer shares of disposable income are quite similar, but the concentration coefficients of taxes are more than twice that of transfers.

As to the finding of no anti-cyclical relationship between the business cycle and the degree of inequality, the decomposition analysis provides several interesting results. Specifically, this analysis indicates that the stability of the income distribution over the economic boom in Norway, lasting from 1993 to 1998, is actually due to two counteracting effects. First, the inequality contributions from earnings and capital income increase during this time period. An increase in the concentration coefficients is the driving force behind capital income becoming more disequalising. Capital income is primarily received by individuals at the upper end of the income distribution, and thus an economic boom leading to increased dividends will lead to more inequality in capital income. By contrast, the increase in the contribution to inequality from earnings is due to a rise in its share of disposable income, largely at the expense of the factor income share of transfers (and self-employment). This is not surprising in light of the strong labour market during this period of time. Second, an increase in the equalising effects of taxes and transfers offset the stronger disequalising effect from earnings and capital income. The increase in the equalising effect of taxes is due to its larger share of disposable income which is directly linked to the increase in the share of earnings in disposable income, since earnings, on average, are taxed more than transfers and income from self-employment. In comparison, the increase in contribution to equality from transfers was driven by a larger concentration coefficient. A plausible explanation is that there is a selection process from welfare to work during an economic boom, which leaves behind the unskilled who are overrepresented

[24] The results presented in Figures A1-A5 in the Appendix demonstrate that the factor decomposition analysis is by and large robust to the choice of inequality measure as well as to the choice of equivalence scale.

at the lower part of the income distribution on welfare rolls. Moreover, there was a substantial shift in government expenditure on social policies away from unemployment benefits to disability benefits during the 1990s. Whilst entitlement to unemployment benefits requires active job search as well as previously earned income through work, disability benefits are granted to those who were permanently out of the labour force due to health problems. Thus, there may have been a change in the population of benefit recipients from the more active part of the population to relatively marginalised subgroups.

As demonstrated by (3.12), each of the three alternative measures of overall inequality depends on inequality within the regions. Thus, an inspection of inequality within regions, and not only the aggregate, may complement the picture of the income distribution in Norway drawn above. Table 1 shows that the degree of inequality does not vary that much between the regions, with a few noticeable exceptions. First and foremost, the capital-region Oslo is distinctive in terms of having the highest income inequality in each year. This is to a large extent due to the fact that capital income has a particular strong disequalising effect in Oslo. Furthermore, Oslo has a relatively large proportion of Non-Western immigrants, who are vastly overrepresented in the lower part of the income distribution.

Table 1. Summary statistics of region-specific inequality measures, 1993-2001

Year	Inequality measure	OECD equivalence scale					EU equivalence scale				
		Mean	Median	Max.	Min.	Mean Deviation from the Mean	Mean	Median	Max.	Min.	Mean Deviation from the Mean
1993	B	0.303	0.300	0.413	0.273	0.013	0.311	0.309	0.419	0.286	0.011
	G	0.207	0.205	0.295	0.183	0.011	0.209	0.207	0.299	0.188	0.010
	C	0.165	0.163	0.244	0.143	0.010	0.165	0.163	0.247	0.146	0.010
1994	B	0.301	0.299	0.408	0.268	0.012	0.310	0.309	0.414	0.282	0.011
	G	0.206	0.203	0.290	0.179	0.011	0.209	0.206	0.294	0.182	0.010
	C	0.164	0.161	0.239	0.139	0.010	0.164	0.161	0.242	0.140	0.010
1995	B	0.299	0.295	0.414	0.270	0.013	0.308	0.305	0.419	0.283	0.012
	G	0.206	0.203	0.297	0.184	0.011	0.208	0.206	0.301	0.187	0.011
	C	0.165	0.162	0.246	0.146	0.011	0.164	0.161	0.249	0.145	0.011
1996	B	0.300	0.296	0.424	0.271	0.013	0.308	0.306	0.430	0.284	0.012
	G	0.207	0.204	0.311	0.183	0.012	0.209	0.206	0.317	0.186	0.011
	C	0.166	0.162	0.261	0.143	0.011	0.165	0.162	0.267	0.143	0.011
1997	B	0.302	0.301	0.425	0.263	0.015	0.311	0.307	0.430	0.276	0.014
	G	0.211	0.209	0.313	0.178	0.014	0.213	0.210	0.317	0.181	0.014
	C	0.170	0.167	0.264	0.140	0.014	0.170	0.166	0.268	0.139	0.014
1998	B	0.294	0.292	0.414	0.258	0.014	0.302	0.299	0.418	0.269	0.014
	G	0.205	0.203	0.303	0.176	0.013	0.207	0.205	0.307	0.178	0.013
	C	0.166	0.163	0.255	0.139	0.013	0.165	0.163	0.258	0.138	0.013
1999	B	0.293	0.290	0.419	0.263	0.016	0.301	0.298	0.424	0.274	0.015
	G	0.206	0.202	0.310	0.176	0.015	0.207	0.204	0.314	0.179	0.015
	C	0.167	0.163	0.262	0.139	0.014	0.166	0.162	0.266	0.139	0.015
2000	B	0.302	0.297	0.449	0.263	0.018	0.309	0.305	0.454	0.272	0.017
	G	0.215	0.210	0.345	0.177	0.017	0.216	0.213	0.349	0.179	0.017
	C	0.176	0.172	0.299	0.140	0.016	0.175	0.171	0.303	0.139	0.017
2001	B	0.291	0.289	0.399	0.261	0.013	0.297	0.294	0.403	0.271	0.013
	G	0.201	0.199	0.286	0.176	0.011	0.202	0.199	0.289	0.178	0.011
	C	0.162	0.159	0.236	0.139	0.011	0.160	0.158	0.239	0.138	0.011

4. Conclusions

While theoretical work on the measurement of income by and large presupposes full comparability of incomes, their empirical counterparts are forced to deal with comparability problems along several dimensions such as time, space, and income recipient. To design and evaluate redistribution programs it is necessary for practitioners to provide an understandable picture of income distributions even when

comparability of incomes is restricted. The widespread use of more or less justifiable equivalence scales - to enable comparison of incomes across individuals belonging to households of varying size and composition - illustrates the strive for interpersonal comparability in empirical assessments of income distributions. Since the conditions of identical prices and qualities on goods as well as uniform consumption habits across regions are usually not fulfilled, the conventional analysis of income inequality based on the distribution of equivalent incomes for the entire population of a country may nevertheless be biased. It is thus important to go beyond simply imposing equivalence scales by also introducing and applying methods for measuring inequality that account for heterogeneity in the spatial dimension.

In this paper, we pursue two strategies to cope with problems of comparability of incomes between regions within a country. To obtain estimates of real income, we transform the observed incomes according to relative housing prices across regions. This is done by estimating regional price indices using the Average Basket version of the Symmetric Star method for making multilateral income comparisons. However, it can be argued that inequality measures based on estimates of real income suffer from a lack of welfare basis for reasons ranging from substitution bias in the price indices to systematic differences across regions in the qualities of goods offered as well as consumption habits. Thus, we also introduce a method that enables us to measure income inequality in a heterogeneous population in a meaningful way when incomes cannot be made adequately comparable between regions. On the basis of the distributions of individual equivalent incomes for a set of subgroups determined by individuals' region of residence, we estimate region-specific inequality measures. By aggregating the inequality measures across regions according to their population shares, we obtain an estimate of overall inequality where the between/across-group inequality term is excluded.

Applying a unique register household panel data set covering the entire resident population of Norway in the period 1993-2001, we find that the levels and trends in overall inequality as well as the inequality contributions of various income factors are robust to whether the income inequality analysis is based on the standard approach, which presupposes a homogeneous population, or the methods developed to cope with comparability problems within a country. Consequently, the conventional method for measuring inequality presupposing no heterogeneity in the spatial dimension can be argued to provide reliable guidelines for economic policy. This is at least the case in the setting of a Scandinavian welfare state characterised by generous cash benefits, a comprehensive public sector, and centralised wage setting which presumably contribute to equalising the distribution of incomes between and across regions.

References

Aaberge, R. (1997): “Interpretation of Changes in Rank-Dependent Measures of Inequality”, *Economics Letters*, **55**, 215-219.

Aaberge, R. (2000): "Characterizations of the Lorenz Curves and Income Distributions", *Social Choice and Welfare*, **17**, 639-653.

Aaberge, R. (2001a): "Sampling Errors and Cross-Country Comparisons of Income Inequality”, *Journal of Income Distribution*, **10**, 75-82.

Aaberge, R. (2001b): "Axiomatic Characterization of the Gini Coefficient and Lorenz Curve Orderings", *Journal of Economic Theory*, **101**, 115-132.

Aaberge, R. (2007): “Gini's Nuclear Familiy", forthcoming in *Journal of Economic Inequality*.

Aaberge, R., A. Bjørklund, M. Jantti, P. Pedersen, N. Smith, and T. Wennemo (2000): "Unemployment Shocks and Income Distribution: How Did the Nordic Countries Fare during Their Crises?", *Scandinavian Journal of Economics*, **102**, 77-99.

Aaberge, R., S. Bjerve, and K. Doksum (2005): "Decomposition of Rank-Dependent Measures of Inequality By Subgroups", *Metron - International Journal of Statistics*, **63**, 493-503.

Anand, S. (1983): *Inequality and Poverty in Malaysia*, Oxford University Press, London.
Atkinson, A. (1992): “Measuring Poverty and Differences in Family Composition”, *Economica*, **59**, 1–16.

Atkinson, A. (1996): “Income Distribution in Europe and the United States”, *Oxford Review of Economic Policy*, **12**, 15-28.

Atkinson, A. and F. Bourguignon (1987): “Income Distributions and Differences in Needs”, in G.R. Feiwel (ed.) *Arrow and the Foundations of Theory of Economic Policy*, Macmillian, London.

Atkinson, A., L, Rainwater, and T. Smeeding (1995): “Income Distribution in OECD Countries”, *Social Policy Studies*, **18**, OECD.

Atkinson, A. and A. Brandolini (2001): “Promises and Pitfalls in the Use of Secondary Data-Sets: Income Inequality in OECD Countries as a Case Study”, *Journal of Economic Literature*, **39**, 771-800.

Berrebi, Z. and J. Silber (1985): “Income Inequality Indices and Deprivation: A Generalisation”, *Quarterly Journal of Economics*, **100**, 807-810.

Bhattacharya, N. and B. Mahalanobis (1967): "Regional Disparities in Household Consumption in India", *JASA*, **62**, 143-161.

Brandolini, A. (2006): “Measurement of Income Distributions in Supranational Entities: The Case of the European Union”, forthcoming in S. Jenkins and J. Micklewright (eds.) *Inequality and Poverty Re-examined*, Oxford, Oxford University Press.

Chambaz, E. and C. Maurin (1998): “On Comparing Heterogeneous Populations: Is There Really a Conflict Between the Pareto Criterion and Inequality Aversion?”, *Review of Income and Wealth*, **44**, 497-513.

Coulter, F., F. Cowell and S. Jenkins (1992): "Differences in Needs and Assessment of Income Distributions", *Bulletin of Economic Research*, **44**, 77-124.

Cowell, F. (1980): "On the Structure of Additive Inequality Measures", *Review of Economic Studies*, **47**, 521-531.

Cowell, F. (1988): "Inequality Decomposition: Three Bad Measures", *Bulletin of Economic Research*, **40**, 309-312.

Cowell, F. (1995): *Measuring Inequality*, LSE Handbooks in Economics, 2nd edition, Prentice Hall, London.

Citro, C. and R. Michael (eds.) (1995): *Measuring Poverty: A New Approach*, Washington DC, National Academy Press.

Dagum, C. (1997): "A New Approach to the Decomposition of the Gini Income Inequality Ratio", *Empirical Economics*, **22**, 515-531.

Das, T. and A. Parikh (1982): "Decomposition of Inequality Measures and a Comparative Analysis", *Empirical Economics*, **7**, 23-48.

Diewert, E. (1976): "Exact and Superlative Index Numbers", *Journal of Econometrics*, **4**, 115-145.

Diewert, E. (1999): "Axiomatic and Economic Approaches to International Comparisons", in A. Heston and R. Lipsey (eds.) *International and Interarea Comparisons of Income, Output, and Prices*, NBER Studies of Income and Wealth, 61, Chicago, University of Chicago Press.

Expert Group on Household Income Statistics (2001): *Final Report and Recommendations*, Ottawa, Canada.

Fei, J., C. Ranis, and S. Kuo (1978): "Growth and Family Distribution of Income by Factor Components", *Quarterly Journal of Economics*, **92**, 17-53.

Fjærli, E. and R. Aaberge (2000): "Tax Reforms, Dividend Policy, and Trends in Income Inequality", Discussion Paper, 284, Statistics Norway, Research Department.

Fleurbaey, M., C. Hagnere and A. Trannoy (2003): "Welfare Comparisons with Bounded Equivalence Scales", *Journal of Economic Theory*, **11**, 309-336.

Foster, J. (1983): "An Axiomatic Characterization of the Theil Measure in Income Inequality", *Journal of Economic Theory*, **31**, 105-121.

Foster, J. (1984): "Normative Measurement: Is Theory Relevant?", *American Economic Review*, **84**, 365-370.

Foster, J. and A. Shneyerov (1999): "A General Class of Additively Decomposable Inequality Measures", *Economic Theory*, **14**, 89-111.

Galloway, T. and M. Mogstad (2006): "Årlig og Kronisk Fattigdom i Norge: En Empirisk Analyse av Perioden 1993-2001", *Rapporter*, 20, Statistics Norway. [In Norwegian]

Hill, R. (1997): “A Taxonomy of Multilateral Methods for Making International Comparisons of Prices and Quantities”, *Review of Income and Wealth*, **43**, 49-69.

Hill, R. (2004): “Constructing Price Indexes across Space and Time: The Case of the European Union”, *American Economic Review*, **95**, 1379-1410.

Jenkins, S. and P. Lambert (1993): “Ranking Income Distributions when Needs Differ”, *Review of Income and Wealth*, **39**, 337-356.

Jolliffe, D. (2006) “Poverty, Prices, and Place: How Sensitive is the Spatial Distribution of Poverty to Cost of Living Adjustments?” *Economic Inquiry*, **44**, 296–310.

Lambert, P. and R. Aronson (1993): "Inequality Decomposition Analysis and the Gini-coefficient Revisited", *Economic Journal*, **103**, 1221-1227.

Lambert, P. and X. Ramos (2002): “Welfare Comparisons: Sequential Procedures for Heterogeneous Populations”, *Economica*, **69**, 549-562.

Kakwani, N. (1980): *Income Inequality and Poverty: Methods of Estimation and Policy Applications*, Oxford University Press, Oxford.

Kakwani, N. (1984): “The Relative Deprivation Curve and its Applications”, *Journal of Business and Economic Statistics*, **2**, 384-405.

Lerman, R. and S. Yitzhaki (1985): "Income Inequality Effects by Income Sources: A New Approach and Applications to the United States", *Review of Economics and Statistics*, **67**, 151-156.

Marris, R. (1984): “Comparing the Incomes of Nations: A Critique of the International Comparison Project”, *Journal of Economic Literature*, **22**, 40-57.

Milanovic, B. (2002): “The True World Income Distribution, 1988 and 1993: First Calculations based on Household Surveys Alone”, *Economic Journal*, **112**, 51-92.

Mogstad, M., A. Langørgen, and R. Aaberge (2007): "Region-specific versus Country-specific Poverty Lines in Analysis of Poverty", *Journal of Economic Inequality*, 5, 115-122.

Neary, P. (2004): “Rationalizing Penn World Table: True Multilateral Indices for International Comparison of Real Income”, *American Economic Review*, **94**, 1411-1428.

Pyatt, G. (1976): "On the Interpretation of and Disaggregation of Gini-coefficients", *Economic Journal*, **86**, 243-255.

Pyatt, G., C. Chen, and J. Fei (1980): "The Distribution of Income by Factor Components", *Quarterly Journal of Economics*, **95**, 451-473.

Rao, V. (1969): "Two Decompositions of the Concentration Ratio", *Journal of the Royal Statistical Society*, **132**, 418-425.

Rogoff, K. (1996): “The Purchasing Power Parity Puzzle”, *Journal of Economic Literature*, **34**, 647-668.

Runciman, W. (1966): *Relative Deprivation and Social Justice*, London, Routledge.

Sala-i-Martin, X. (2006): “The World Distribution of Income”, *Quarterly Journal of Economics*, **121**, 351-397.

Schultz, T. (1998): “Inequality in the Distribution of Personal Income in the World: How it is changing and why?”, *Journal of Population Economics*, **11**, 327-368.

Sen, A. (1984): “Poor, Relatively Speaking”, in A. Sen (ed.) *Resources, Values and Development*, First Edition, Oxford, Basil Blackwell Publisher Ltd.

Sen, A. and J. Foster (1997): *On Economic Inequality: Expanded Edition*, Clarendon Press, Oxford.

Silber, J. (1989): "Factor Components, Population Subgroups and the Computation of Gini Index of Inequality", *Review of Economics and Statistics*, **71**, 107-115.

Silber, J. (1993): "Inequality Decomposition by Income Source: A Note", *Review of Income and Statistics*, **75**, 545-547.

Shorrocks, A. (1980): "The Class of Additively Decomposable Inequality Measures", *Econometrica*, **48**, 613-625.

Shorrocks, A. (1984): "Inequality Decomposition by Population Subgroup", *Econometrica*, **52**, 1369-1385.

Short, K., M. Shea, D. Johnson, and T. Garner (1998): “Poverty-Measurement Using the Consumer Expenditure Survey and the Survey of Income and Program Participation”, *American Economic Review*, **88**, 352-356.

Short, K. (2001): “Where We Live: Geographical Differences in Poverty Thresholds”, Poverty Measurement Working Paper, US Census Bureau.

Smeeding, T., P. Saunders, J. Coder, S. Jenkins, J. Fritzell, A. Hagenaars, R. Hauser, and M. Wolfson (1993): “Poverty, Inequality, and Family Living Standards: The Effect of Non-cash Subsidies”, *Review of Income and Wealth*, **39**, 229-256.

Statistics Norway (2000): “Classification of Economic Regions”, Official Statistics of Norway, C634, Statistics Norway.

Statistics Norway (2002): “Konsumprisindeksen 1995-2000”, Official Statistics of Norway, C680, Statistics Norway. [In Norwegian]

Wiles, P. (1974): *Distribution of Income: East and West.* North Holland, Amsterdam.

Yao, S. (1997): "Decomposition of Gini coefficients by Income Factors: A New Approach and Application", *Applied Economics Letters*, **4**, 27-31.

Yitzhaki, S. (1979): “Relative Deprivation and the Gini-coefficient”, *Quarterly Journal of Economics*, **32**, 321-324.

Yitzhaki, S. (1994): "Economic Distance and Overlapping of Distribution", *Journal of Economics*, **61**, 147-159.

Yitzhaki, S. and R. Lerman (1991): "Income Stratification and Income Inequality", *Review of Income and Wealth*, **37**, 313-329.

Yotopoulos, P. (1989): “Distribution of Real Income: Within Countries and By World Income Classes”, *Review of Income and Wealth*, **35**, 357-376.

Figure A1. Contribution by transfers to inequality in disposable income when country-specific and region-specific inequality measures based on observed and real incomes form the basis, 1993-2001

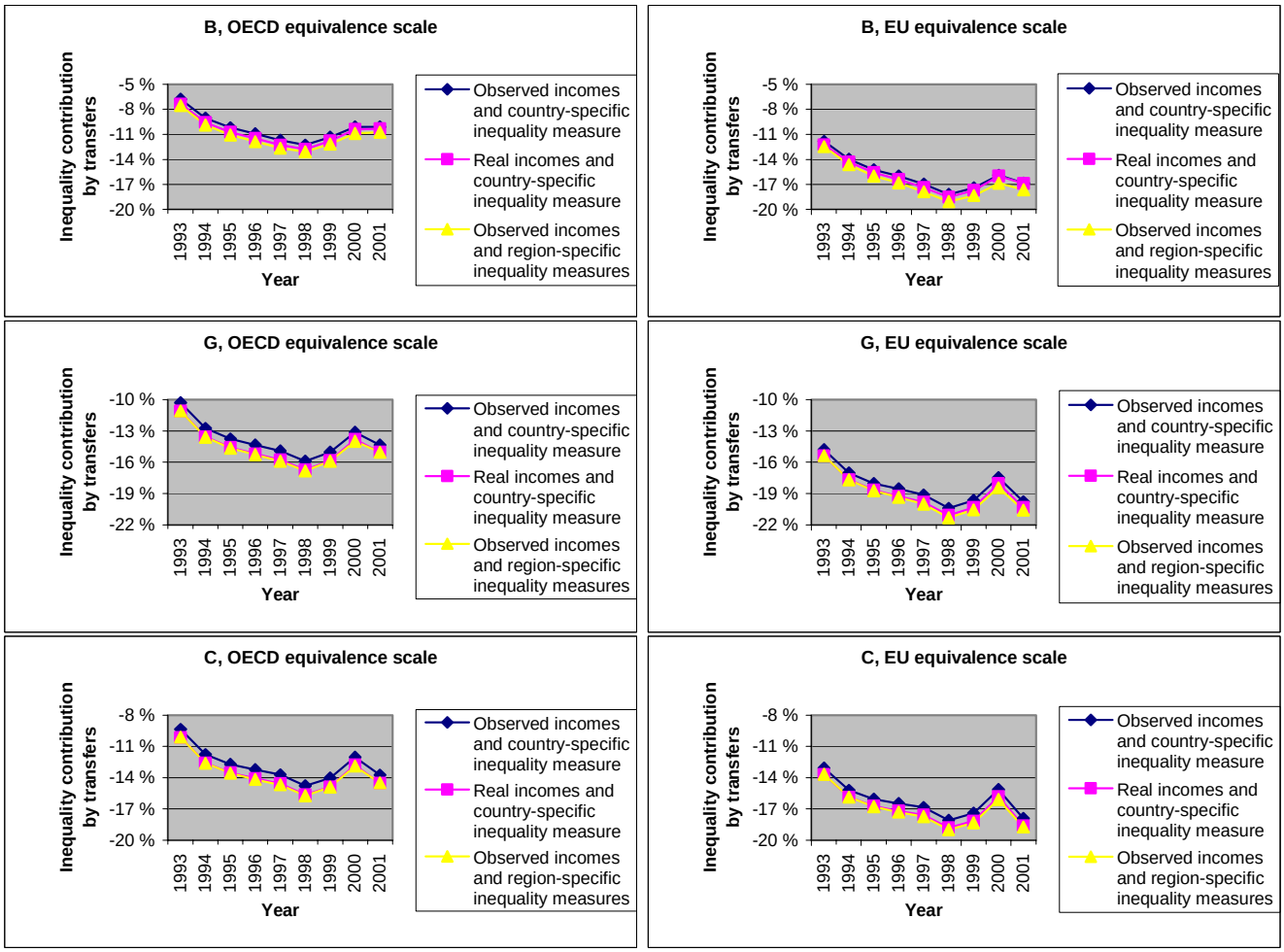


Figure A2. Contribution by capital income to inequality in disposable income when country-specific and region-specific inequality measures based on observed and real incomes form the basis, 1993-2001

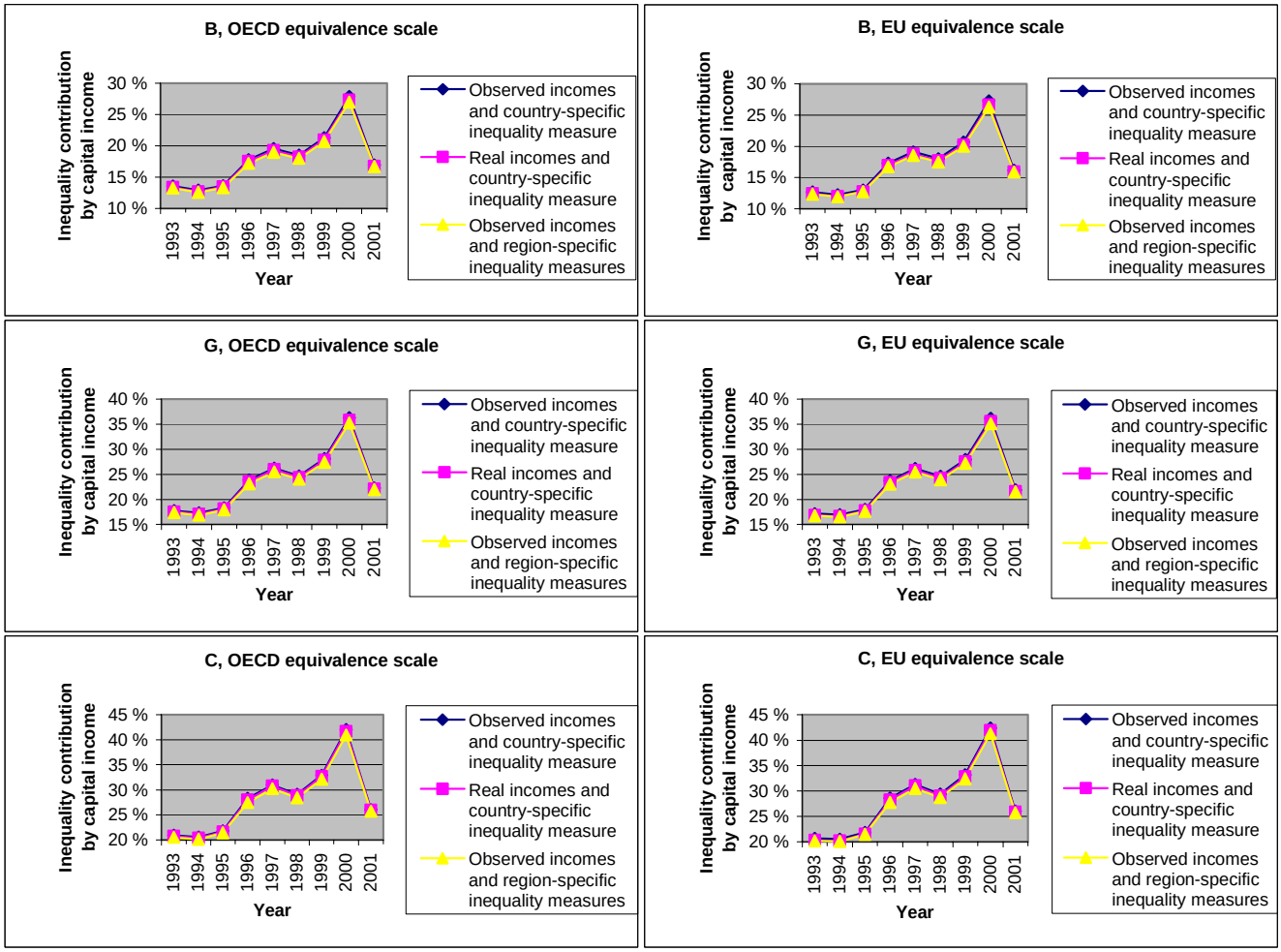


Figure A3. Contribution by tax to inequality in disposable income when country-specific and region-specific inequality measures based on observed and real incomes form the basis, 1993-2001

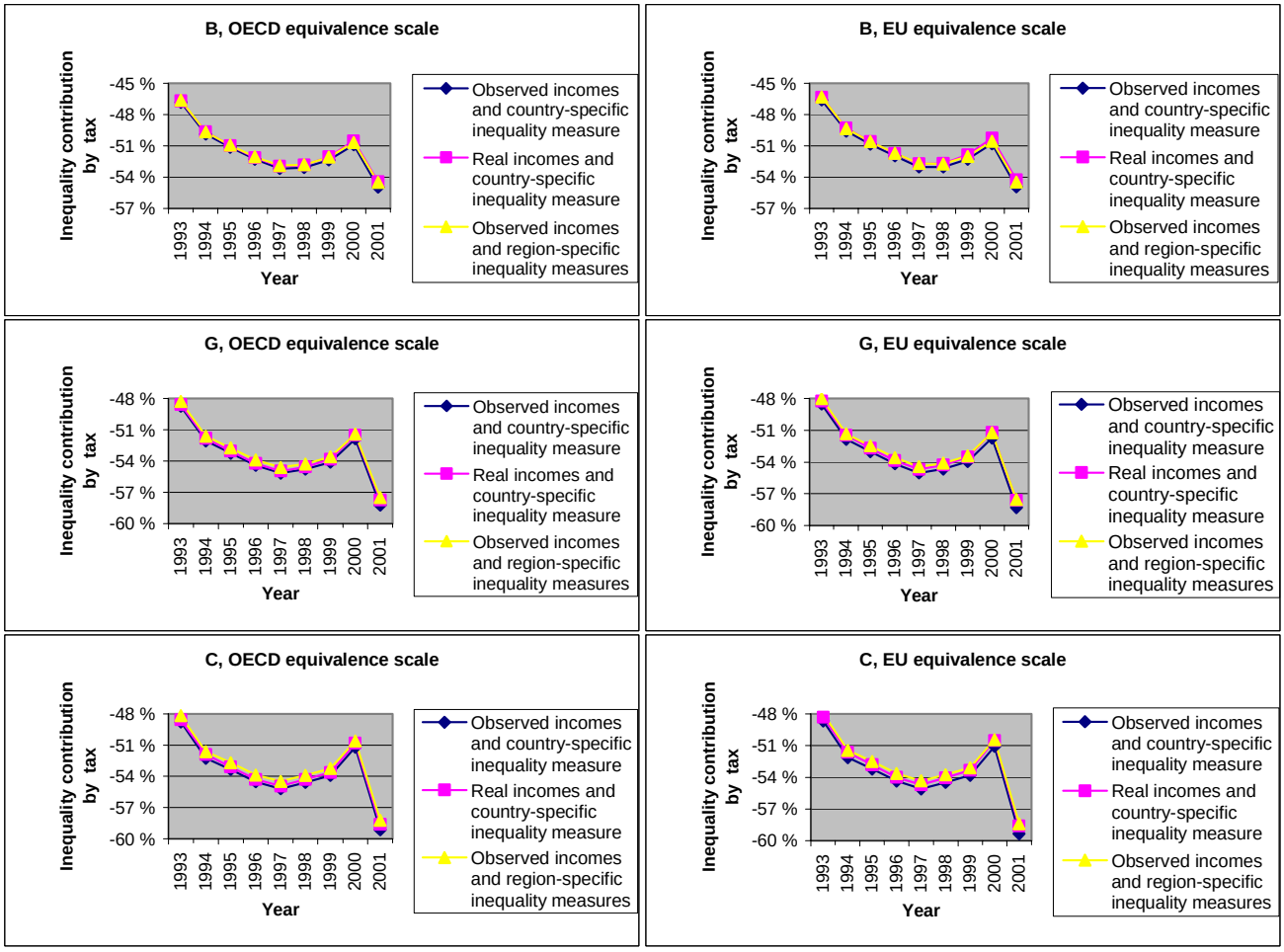


Figure A4. Contribution by self-employment income to inequality in disposable income when country-specific and region-specific inequality measures based on observed and real incomes form the basis, 1993-2001

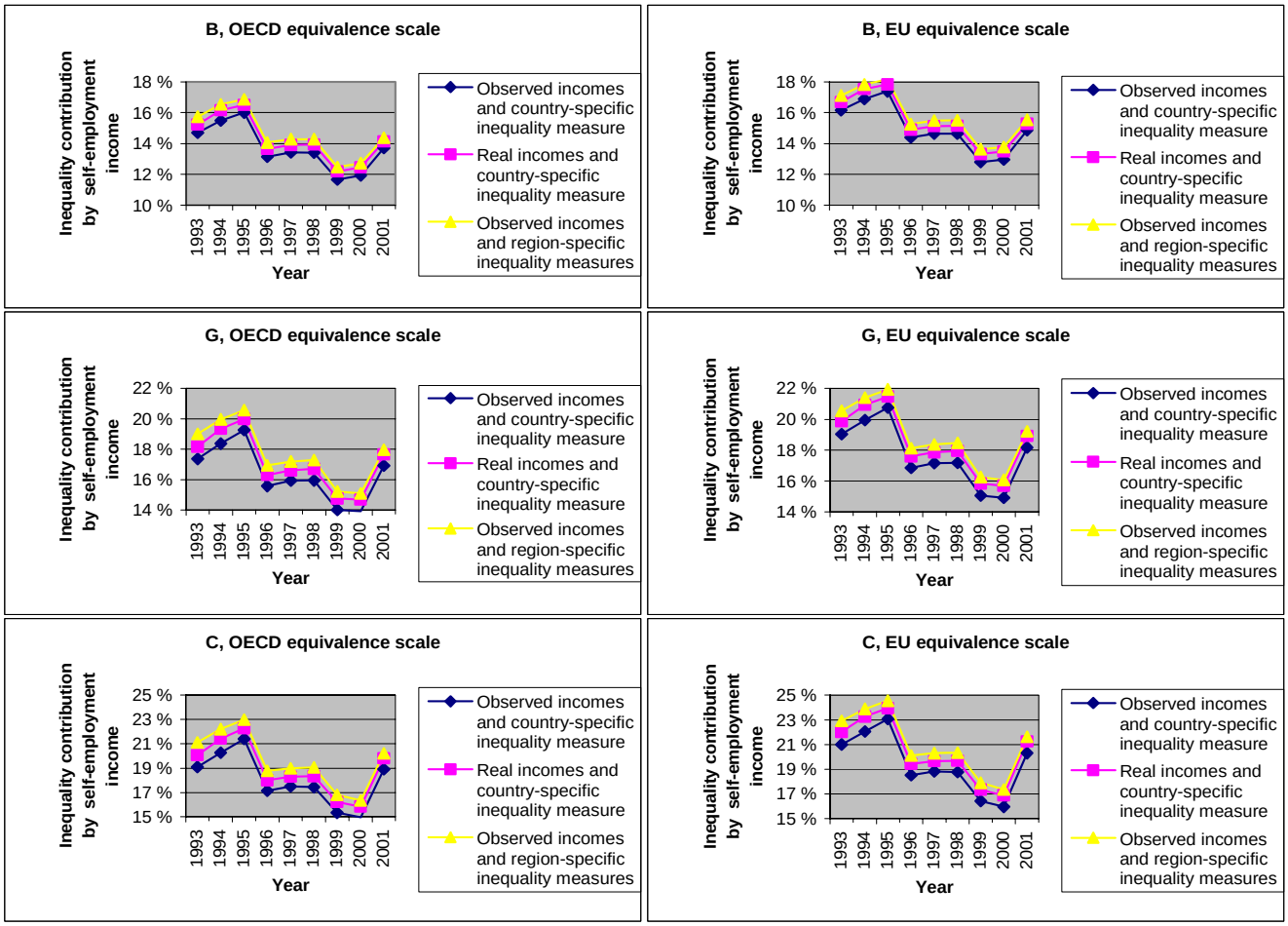


Figure A5. Contribution by earnings to inequality in disposable income when country-specific and region-specific inequality measures based on observed and real incomes form the basis, 1993-2001