# Clarifying Poverty Decomposition

Adrian Muller*†

WORKING PAPER - THIS VERSION: 17.11.2008

**Abstract:** I discuss how poverty decomposition methods relate to integral approximation, which is the foundation of decomposition of the temporal change of a quantity into key drivers. This offers a common framework for the different decomposition methods used in the literature, clarifies their often somewhat unclear theoretical underpinning and identifies the methods' shortcomings. In light of integral approximation, many methods actually lack a sound theoretical basis and they usually have an ad-hoc character in assigning the residual terms to the different key effects. I illustrate these claims for the Shapley-value decomposition and methods related to the Datt-Ravallion approach and point out difficulties in axiomatic approaches to poverty decomposition. Recent developments in energy and pollutant decomposition offer

*Socio-economic Institute, University of Zürich, Blümlisalpstrasse 10, 8006 Zürich, Switzerland; phone: 0041-44-634 37 35; e-mail: adrian.mueller@soi.uzh.ch

†Many thanks to Tony Shorrocks for very helpful remarks. Many thanks for helpful remarks and inspiring discussions to an anonymous referee and to Dilip Mookherjee, to Uma Rani, Åsa Löfgren and workshop participants at the UNU-WIDER conference on Frontiers of Poverty Analysis 2008 in Helsinki and at the Annual Conference of the Verein für Socialpolitik, Research Committee Development Economics, 2008 in Zurich. The usual disclaimer applies.

some promising methods, but ultimately, further development of poverty decomposition should account for the basis in integral approximation.

**Keywords:** poverty analysis, poverty measures, decomposition, Shapley-value, inequality

**JEL:** I32, C43

# 1 Introduction

Decomposing some key variable in several components to better understand the key variable is a common exercise in many areas. Classical in poverty analysis are the (static) identification of the contribution of different population groups, population characteristics or income types to overall inequality at a given point of time (see e.g. Shorrocks 1982, Foster et al. 1984, Shorrocks 1999, Morduch and Sicular 2002, Borooah 2005, Kolenikov and Shorrocks 2005), the (dynamic) identification of the contributions of different driving forces to the evolution of poverty and inequality measures over time (e.g. Shorrocks 1999; Kakwani 2000; for a recent review, see Heshmati 2004) or the identification of contributions in a poverty measure referring to transient and chronic poverty, respectively (e.g. Duclos et al. 2006; on these three types of decomposition, see also Shorrocks 2008). Decomposition[1] plays

[1] "Decomposition" refers to the particular type of methods to identify key drivers of the evolution of some variable of interest that are the topic of this paper and will be explained below. For completeness, I mention that regression analysis clearly also delivers such identification. Even more, regression analysis, being a statistical technique, allows for statistical inference. Decomposition, on the other hand, is purely descriptive. However, it complements regression analysis, as it can give a detailed picture of past developments and as it has much lower data requirements. It can, for example, help to identify structural

an important role in many more areas besides poverty analysis. Examples are the decomposition of the temporal evolution of energy use and pollutant emissions into key drivers (e.g. Ang 1995, 2004; Bruvoll and Larsen 2004), the so-called "growth accounting" to investigate economic growth (e.g. Barro and Sala-I-Martin 2003) and the general index number theory as developed for price and quantity indices (e.g. Diewert and Nakamura 1993).

In this paper, I am mainly concerned with "dynamic" decomposition, that is the second type of decomposition mentioned above. A dynamic approach can also be differentiated to account for the effects of group structures and characteristics (see e.g. Ang 1995, 2004; see also footnote 3).[2] Dynamic decomposition is based on integral approximation and the price and quantity index literature, for example, is partly aware of this (Trivedi 1981, Balk 2005). The awareness of this basis in integral approximation has, however, been lost in the literature on poverty decomposition. Due to the lack of this connection to the underlying basic formalism, current efforts to develop optimal decomposition approaches often seem somewhat arbitrary. This is for example the case for the often-used Shapley-value decomposition (Shorrocks 1999; Baye 2005).

Explicit reference to integral approximation as the underlying formalism of decomposition offers a common framework for the decomposition approaches most frequently applied in the poverty and inequality context, i.e.

---

breaks or other patterns in the data that should be accounted for when setting up a regression analysis.

[2]The "decomposability" of "decomposable" poverty indices refers to certain advantageous criteria with respect to such group structure and the static decomposition into contributions of different groups and is not linked to the dynamic decomposition I am concerned with here.

the Shapley-value based decomposition and the decomposition methods similar to the one presented in Datt and Ravallion (1992). I will show that these methods are special and not always consistent approaches to approximate the underlying integrals. In general, appreciating this common ground in integral approximation could help to develop improved decomposition methods.

Reference to the basis in integral approximation would also shed a new light on the discussion of the residual in decomposition. The residual is present in some classical approaches to poverty decomposition and usually given the somewhat vague interpretation of interaction effects (e.g. Datt and Ravallion 1992). This interpretation is often criticized and the absence of a residual term in the newer approaches related to the Shapley-value is seen as an advantage (Baye 2005). A zero residual is however not necessarily a good criterion to identify optimal decomposition methods. If tied to the underlying integral approximations, the presence of some residual due to approximation errors is natural.

Section 2 introduces the general formalism of decomposition and illustrates how it is linked to integral approximation. Section 3 presents some of the main methods of poverty decomposition currently applied and illustrates how they relate to each other and to the general formalism based on integral approximation. Section 4 concludes.

## 2 A General Formalism for Decomposition

The aim of dynamic decomposition analysis is to identify and assess the (relative) magnitude of different variables driving the time development of a key quantity. One example is how much changes in mean income and the income distribution contribute to changes in total poverty within a country.

I develop the following general formalism. I will show in section 3 how the methods commonly used for poverty decomposition can be seen as special cases of this general formalism. The key quantity of interest shall be $P(t)$ $=P(x_1(t), ..., x_m(t))$, e.g. a poverty measure (examples are the FGT index, Foster et al. 1984, or the Watts index, Chakravarty et al. 2008), depending on $m$ time-dependent variables $x_i(t), i = 1, ...m, t \in [T_0, T_n]$ (e.g. the head-count or income-gap ratio, the poverty line, the mean and variance of the income distribution, etc.).[3]

Considering this poverty measure, we have the change in $P$ from $t = T_0$ to $T_n$: $\Delta P_{T_0,T_n} := P(T_n) - P(T_0)$. What we want to have is a formula of the following form: $\Delta P_{T_0,T_n} = \sum_{i=1}^{m} f_i(x_i(T_n), x_i(T_0))$, where the overall temporal change in $P$ is decomposed into a sum of contributions $f_i$ that depend on the temporal changes in the respective driver variable $x_i$ only.

I make a short detour to illustrate the task at hand with a simple example: consider a cuboid with sides $x_1$, $x_2$ and $x_3$ that changes volume over time, from $V(T_0) = x_1(T_0)x_2(T_0)x_3(T_0)$ at $t = T_0$ to $V(T_n) = x_1(T_n)x_2(T_n)x_3(T_n)$ at $t = T_n$. The task of decomposition is to assign the difference in volume $\Delta V = V(T_n) - V(T_0)$ to differences $\Delta x_i$ in the single edges $x_1$, $x_2$ and $x_3$ (see figure 1).

[3] $P$ can further be differentiated according to some group-structure of interest, i.e. $P = \sum_{g=1}^{G} P^g$, where $P^g$ is the value for $P$ referring to group $g$. Depending on the decomposition method applied, effects of temporal changes in group composition and also relative comparison of the development of different subgroups can be assessed. Much on this can be found in energy and pollution decomposition analysis (see e.g. Ang 2004). Including group structure does not change the general argument and I thus use the simpler notation without this additional structure.

Insert Figure 1 here

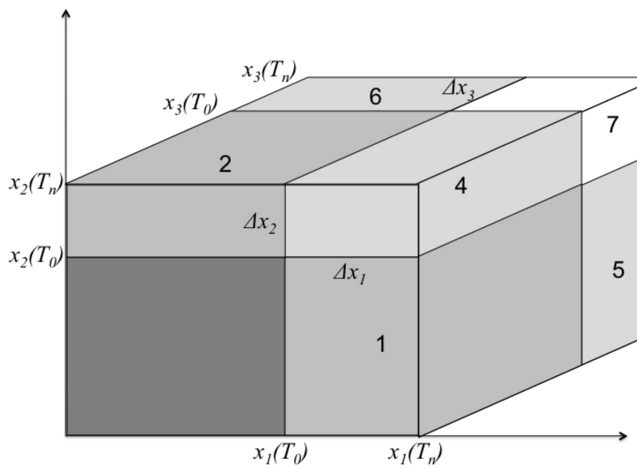


Figure 1: Illustration of the basic task in decomposition

It is natural to assign cuboid 1 to changes in $x_1$ (while $x_2$ and $x_3$ stay constant), cuboid 2 to changes in $x_2$ ($x_1$ and $x_3$ staying constant) and cuboid 3 (not visible in the picture) to changes in $x_3$ ($x_1$ and $x_2$ staying constant). It is, however, not a priori clear how to assign cuboid 4 to changes in $x_1$ and $x_2$ (while $x_3$ stays constant), and correspondingly for cuboids 4 and 5. It is neither clear how to assign cuboid 7 to changes in $x_1$, $x_2$ and $x_3$.[4]

I now come back to the general problem. A natural candidate for such an assignment of driver-specific contributions (a decomposition into separate contributions from each driver variable) is found by the following considerations. First, consider infinitesimal changes of $P$, using the chain rule: $\frac{\mathrm{d}P}{\mathrm{d}t} = \sum_{i=1}^{m} \frac{\partial P}{\partial x_i} \frac{\partial x_i}{\partial t}$. If only $x_{i'}$ changes with $t$ and all other partial derivatives

[4]Setting $P(x_1(t), x_2(t), x_3(t)) = x_1(t)x_2(t)x_3(t)$ ties this example to the general discussion in this section.

$\frac{\partial x_i}{\partial t}$ for $i \neq i'$ are zero, we have $\frac{\mathrm{d}P}{\mathrm{d}t} = \frac{\partial P}{\partial x_{i'}}\frac{\partial x_{i'}}{\partial t}$. $\frac{\partial P}{\partial x_{i'}}\frac{\partial x_{i'}}{\partial t}$ thus captures the contribution of changes in $x_{i'}$ to changes in $P$ when $t$ changes infinitesimally, assuming the other drivers $x_{i\neq i'}$ do not change. It is thus natural to assign the term $\frac{\partial P}{\partial x_i}\frac{\partial x_i}{\partial t}$ to the contribution of changes in $x_i$ to overall changes in $P$ when $t$ changes infinitesimally.

We now consider non-infinitesimal changes in $t$, say from $t = T_0$ to $T_n$. $\Delta P_{T_0,T_n}$ can by definition be written as the integral from $t = T_0$ to $t = T_n$ of the derivative of $P$:

$$\Delta P_{T_0,T_n} \quad := \quad P(T_n) - P(T_0) = \int_{T_0}^{T_n} \frac{\mathrm{d}P}{\mathrm{d}t}\mathrm{d}t. \tag{1}$$

Inserting $\frac{\mathrm{d}P}{\mathrm{d}t} = \sum_{i=1}^{m} \frac{\partial P}{\partial x_i}\frac{\partial x_i}{\partial t}$ in equation (1) gives

$$\begin{aligned}\Delta P_{T_0,T_n} \quad &= \quad \int_{T_0}^{T_n} \frac{\mathrm{d}P}{\mathrm{d}t}\mathrm{d}t = \int_{T_0}^{T_n} \Big(\frac{\partial P}{\partial x_1}\frac{\partial x_1}{\partial t} + \frac{\partial P}{\partial x_2}\frac{\partial x_2}{\partial t} + ... + \frac{\partial P}{\partial x_m}\frac{\partial x_m}{\partial t}\Big)\mathrm{d}t = \\ &= \quad \int_{T_0}^{T_n} \frac{\partial P}{\partial x_1}\frac{\partial x_1}{\partial t}\mathrm{d}t + \int_{T_0}^{T_n} \frac{\partial P}{\partial x_2}\frac{\partial x_2}{\partial t}\mathrm{d}t + ... + \int_{T_0}^{T_n} \frac{\partial P}{\partial x_m}\frac{\partial x_m}{\partial t}\mathrm{d}t. \end{aligned} \tag{2}$$

Again, *ceteris paribus*-assessment is illustrative: assuming that only $x_{i'}$ changes with $t$, only the corresponding term $\Delta P^{x_{i'}}_{T_0,T_n} := \int_{T_0}^{T_n} \frac{\partial P}{\partial x_{i'}}\frac{\partial x_{i'}}{\partial t}\mathrm{d}t$ survives. Inspired by this and by the preceding discussion of infinitesimal changes, the part containing the derivative with respect to $x_{i'}$, i.e. $\Delta P^{x_{i'}}_{T_0,T_n}$, is then interpreted as the contribution from $x_{i'}$ to changes in $P$ when $t$ changes from $T_0$ to $T_n$ and all other driver variables $x_i$, $i \neq i'$ are kept constant.

Usually, the functions involved are not known for all points $t \in [T_0, T_n]$, but only for some discrete points of time, most often equally spaced (e.g. annually): $T_0, T_1, T_2, ..., T_{n-1}, T_n$. We can thus write $\Delta P^{x_i}_{T_0,T_n} := \int_{T_0}^{T_n} \frac{\partial P}{\partial x_i}\frac{\partial x_i}{\partial t}\mathrm{d}t = \int_{T_0}^{T_1} \frac{\partial P}{\partial x_i}\frac{\partial x_i}{\partial t}\mathrm{d}t + \int_{T_1}^{T_2} \frac{\partial P}{\partial x_i}\frac{\partial x_i}{\partial t}\mathrm{d}t + ... + \int_{T_{n-1}}^{T_n} \frac{\partial P}{\partial x_i}\frac{\partial x_i}{\partial t}\mathrm{d}t = \sum_{k=1}^{n} \Delta P^{x_i}_{T_{k-1},T_k}$.

Thus, to calculate the expressions for the contributions of the various variables $x_i$ $(i = 1, ..., m)$ in equation (2), integrals of the following form have to be calculated[5]:

$$\Delta P^{x_i}_{T,T+1} = \int_T^{T+1} \frac{\partial P(x_1, ..., x_m)}{\partial x_i} \frac{\partial x_i}{\partial t} \mathrm{d}t. \tag{3}$$

Hereby, the integrands are basically known at the endpoints only.

Decomposing $P$ thus boils down to solving such integrals. Because of the lack of information, though, i.e. the lack of knowledge on the underlying functions besides for the boundary values $T$ and $T+1$, this is essentially an approximation problem. The integral has to be approximated by the values of the integrand at the endpoints of the integration range. In addition, the presence of derivatives cause problems, as for $x_i(t), i = 1, ..., m$, only the values of the functions but not of the derivatives are known for the endpoints. In this case, some approximation of the derivatives is necessary as well. As the poverty measure $P$ is known as a function of its variables $x_i$, the derivatives of $P$ with respect to $x_i$ are known. The integral can thus be written as a function $J$ or $\tilde{J}$ of the values at the end-points:[6]

[5]From now on, to keep notation simple, $T$ stands for any value $T_k$, $k = 1, ..., n-1$, and $T+1$ correspondingly stands for $T_{k+1}$.

[6]$J$ includes the derivatives of $x_i$ directly, while they are approximated in $\tilde{J}$. For $\tilde{J}$, I chose the general formulation including $x_i(T-1)$ and $x_i(T+2)$, as they may enter the formula depending on how the derivatives are approximated. An approximation of the derivative at $T+1$ from the right side, for example, usually depends on the value at $T+2$.

$$
\begin{aligned}
&\Delta P^{x_i}_{T,T+1} \approx \\
\approx\ & J\Big(P(T), x_i(T), \frac{\partial P}{\partial x_i}(T), \frac{\partial x_i}{\partial t}(T), \\
& P(T+1), x_i(T+1), \frac{\partial P}{\partial x_i}(T+1), \frac{\partial x_i}{\partial t}(T+1)\Big) \approx \\
\approx\ & \tilde{J}\Big(P(T), x_i(T), P(T+1), x_i(T+1), \\
& \frac{\partial P}{\partial x_i}(T), \frac{\partial P}{\partial x_i}(T+1), x_i(T-1), x_i(T+2)\Big),
\end{aligned} \tag{4}
$$

The simplest approximations of $\Delta P^{x_i}_{T,T+1}$ are based on replacing the true function by different types of step-functions. Thus, the integral is replaced with the product of the value of the integrand at the upper or lower end-point times the distance on the ordinate $\Delta T$, in this case equaling one: $J = \frac{\partial P}{\partial x}\frac{\partial x}{\partial t}|_{T+1 \text{ resp. } T}$. A related approach is to replace the integral by the trapezoid given by joining the upper and lower end-point with a straight line (see figure 2). This corresponds to the average of the two previous approaches: $J = [\frac{\partial P}{\partial x}\frac{\partial x}{\partial t}(T+1) + \frac{\partial P}{\partial x}\frac{\partial x}{\partial t}(T)]/2$.[7] These three approximations are analogous to classical indices in the price/quantity context (the Laspeyres, Paasche and Marshall-Edgeworth index[8]). Especially the first two have sev-

[7] $\frac{\partial x}{\partial t}(T+1)$ and $\frac{\partial x}{\partial t}(T)$ are then usually approximated by the slope of the straight line joining the endpoints, i.e. $\frac{\partial x}{\partial t}(T+1) \approx x(T+1) - x(T) \approx \frac{\partial x}{\partial t}(T)$, thus giving the same value and simplifying decomposition formulae. This strategy could be criticized because of its inconsistency by taking the approximation from the right for the value at the left boundary $T$ and the value from the left at $T+1$. This leads to potentially different results for $\frac{\partial x}{\partial t}(T)$ depending on whether it is part of a term between $T-1$ and $T$ or between $T$ and $T+1$. However, the strategy makes sense if seen in the context of replacing the whole unknown function with straight lines joining the known values, as in the method just mentioned.

[8] In the light of the illustrative example given in figure 1, the Laspeyres approach assigns

eral disadvantages, though, such as a usually rather large residual term or the asymmetry regarding the boundaries (Ang 2004).

Insert Figure 2 here

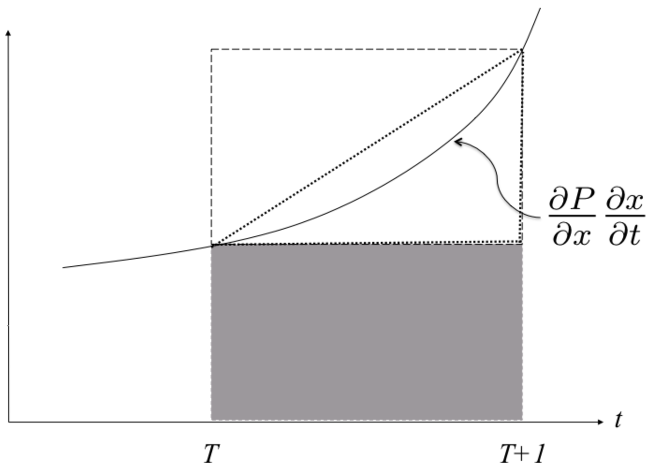


Figure 2: Step-functions for integral approximation. The grey rectangle is the Paasche method, this plus the dashed rectangle give Laspeyres, the grey rectangle and the dotted triangle Marshall-Edgeworth

cuboids 1, 4, 5 and 7 to changes in $x_1$, 2, 4, 6 and 7 to changes in $x_2$ and 3 (not visible), 5, 6 and 7 to changes in $x_3$. This leads to a negative residual of $-(1+2+3+2*(4+5+6)+3*7)$; Paasche assigns cuboid 1 to changes in $x_1$, 2 to changes in $x_2$ and 3 (not visible) to changes in $x_3$, leading to a positive residual of $4+5+6+7$; Marshall-Edgeworth assigns 1 and half of $(4+5)$ and a third of 7 to changes in $x_1$ and correspondingly for $x_2$ and $x_3$. It thus has a zero residual in this example.

# 3 Poverty Decomposition

Poverty decomposition usually refers to decomposing some kind of poverty measure $P$, often the classical measure introduced in Foster et al. (1984), into parts corresponding to the effects of temporal changes in the mean income $\mu$, the income distribution $L$ and the poverty line $z$: $P = P(\mu, L, z)$.[9] This can be normalized by $z$, i.e. the function to be investigated afterward depends on only two instead of three variables: $\bar{P}(\frac{\mu}{z}, \frac{L}{z})$.

There is a range of different methods for poverty decomposition. The choice of a certain method is sometimes based on some formal symmetry arguments or axioms (Shorrocks 1982; Tsui 1996; Kakwani 2000), but in most cases it is rather ad-hoc. There is no awareness of the underlying approximation problem, although the decomposition methods proposed can be understood in this frame (see section 3.3 below). It follows from the discussion above that the general decomposition of the poverty measure reads

$$\Delta P_{T,T+1} = \int_T^{T+1} \frac{\partial P}{\partial \mu}\frac{\partial \mu}{\partial t}\mathrm{d}t + \int_T^{T+1} \frac{\partial P}{\partial L}\frac{\partial L}{\partial t}\mathrm{d}t + \int_T^{T+1} \frac{\partial P}{\partial z}\frac{\partial z}{\partial t}\mathrm{d}t, \qquad (5)$$

where the integrals involved have the same structure as discussed above and similar problems related to their approximation arise.

In the following, I introduce the most common methods for decomposition of changes in poverty or inequality measures (the decompositions in the spirit of Datt and Ravallion (1992), the Shapley-value decomposition and

[9] The income distribution can be characterized by one or more variables, most commonly by its moments (mean, variance, skewness, kurtosis, etc.); in the following, I always use one variable only to characterise it, $L$, as more variables can be treated analogously.

some further related approaches) and show how they relate to the general framework presented above.

### 3.1 Common Approaches to Poverty Decomposition

Most poverty decomposition approaches assume that the contribution of one variable to total change in poverty can be separated if all other variables are kept constant, i.e. if an unobserved "counterfactual situation" is correctly constructed. In particular, the choice of the time period, in which to keep the other variables constant, is crucial and various possibilities for this differentiate the methods. This approach leads to decompositions such as (taking the normalized form with $\bar{\mu} := \frac{\mu}{z}$ and $\bar{L} := \frac{L}{z}$)

$$
\begin{aligned}
\Delta \bar{P}_{T,T+1} &= \bar{P}(\bar{\mu}(T+1), \bar{L}(T+1)) - \bar{P}(\bar{\mu}(T), \bar{L}(T)) = \\
&= \Big[\bar{P}(\bar{\mu}(T+1), \bar{L}(T+1)) - \bar{P}(\bar{\mu}(T), \bar{L}(T+1))\Big] + \\
&\quad + \Big[\bar{P}(\bar{\mu}(T+1), \bar{L}(T+1)) - \bar{P}(\bar{\mu}(T+1), \bar{L}(T))\Big] + \bar{R} = \\
&= \bar{\mu}(\text{i.e. growth})\text{-effect} + \bar{L}(\text{i.e. inequality})\text{-effect} + \bar{R},
\end{aligned}
\tag{6}
$$

where $\bar{R}$ is the residual - also referred to as the interaction effect between growth and changes in inequality, given by $\bar{R} = \bar{P}(\bar{\mu}(T), \bar{L}(T+1)) - \bar{P}(\bar{\mu}(T+1), \bar{L}(T+1)) + \bar{P}(\bar{\mu}(T+1), \bar{L}(T)) - \bar{P}(\bar{\mu}(T), \bar{L}(T))$ (Datt and Ravallion 1992; Baye 2004).

The residual has a similar structure as the decomposition itself and the whole formula has a rather ad-hoc character by adding and subtracting terms to get the effects of interest and then correcting for this by collecting their corresponding negatives in the residual, thus guaranteeing the validity of the formula. Datt and Ravallion (1992) also observe that this residual can be quite large, thus invalidating the whole approach. Equation (6) also depends on the period chosen as base period, as it is not symmetric in $T$ and $T+1$. This method nevertheless is applied without discussion of potential problems, e.g. in Grootaert (1995) or Kraay (2006).

A similar approach is proposed by Jain and Tendulkar (1990),

$$\begin{aligned}\Delta\bar{P}_{T,T+1} &= \Big[\bar{P}(\bar{\mu}(T+1),\bar{L}(T+1)) - \bar{P}(\bar{\mu}(T),\bar{L}(T+1))\Big] + \\ &\quad + \Big[\bar{P}(\bar{\mu}(T),\bar{L}(T+1)) - \bar{P}(\bar{\mu}(T),\bar{L}(T))\Big], \qquad (7)\end{aligned}$$

where the residual is zero, but the two effects are calculated with reference to different base periods and the decomposition is again not symmetric in $T$ and $T+1$. For completeness, I mention that also the (generalized) Oaxaca-Blinder decomposition is of a similar spirit, as can be seen from equation (2) in Yun (2004).

This situation led other authors (e.g. Kakwani 2000; Mazumdar and Son 2001; Bhanumurthy and Mitra 2003; Son 2003) to suggest a symmetric alternative of this decomposition by averaging the formulae with base periods $T$ and $T+1$. Kakwani (2000) in particular motivates this by proposing a set of axioms any poverty decomposition should fulfill (cf. footnote 14 below). This leads to a symmetric decomposition without residual and the growth and inequality effects have the same combination of mixed base periods:

$$\Delta\bar{P}_{T,T+1} = \frac{1}{2}\Big[\bar{P}(\bar{\mu}(T+1),\bar{L}(T+1)) - \bar{P}(\bar{\mu}(T),\bar{L}(T+1)) + \\ +\bar{P}(\bar{\mu}(T+1),\bar{L}(T)) - \bar{P}(\bar{\mu}(T),\bar{L}(T)\Big] + \\ + \frac{1}{2}\Big[\bar{P}(\bar{\mu}(T+1),\bar{L}(T+1)) - \bar{P}(\bar{\mu}(T+1),\bar{L}(T)) + \\ +\bar{P}(\bar{\mu}(T),\bar{L}(T+1)) - \bar{P}(\bar{\mu}(T),\bar{L}(T))\Big]. \quad (8)$$

## 3.2 Generalizations and Further Developments

Decomposition (8) can be generalized to any numbers of variables. Given a poverty measure $P$ depending on $m$ variables $x_1, ..., x_m$, the contribution of $x_i$ to changes in $P$ can be defined to be a combination of all terms of the following form,

$$\Delta P^{x_i}_{T,T+1}(\pi_{s-1,m-s}) = [P(..., x_i(T+1), ...) - P(..., x_i(T), ...)], \quad (9)$$

where all other variables than $x_i$ are evaluated at either $T+1$ or $T$ in both terms to the right and $x_i$ is evaluated at $T+1$ in the first and at $T$ in the second term. This is captured by $\pi_{s-1,m-s}$, which is any $m-1$-vector with $s-1$ entries $T+1$ and $m-s$ entries $T$. The elements of this vector indicate at which time the variables other than $x_i$, i.e. $x_1, ..., x_{i-1}, x_{i+1}, ..., x_m$, are taken in both the terms on the right hand side in equation (9).[10] For $m$ variables, a certain combination of $s$ variables taken at $T+1$ and $m-s$

[10] This is a type of *ceteris paribus* reasoning employing all combinations of how the other variables can stay constant: each at $T$ or at $T+1$. For illustration, I give some of the terms for $i=5$: $\Delta P^{x_1}_{T,T+1}(T+1,T,T,T) = [P(x_1(T+1), x_2(T+1), x_3(T), x_4(T), x_5(T)) - P(x_1(T), x_2(T+1), x_3(T), x_4(T), x_5(T))]$; $\Delta P^{x_3}_{T,T+1}(T, T+1, T, T+1) = [P(x_1(T), x_2(T+1), x_3(T+1), x_4(T), x_5(T+1)) - P(x_1(T), x_2(T+1), x_3(T), x_4(T), x_5(T+1))]$.

at $T$ thus shows up in the final expression $s$ times with a positive sign, stemming from the positive part of equation (9), for each of the $s$ variables at $T+1$. And correspondingly, it shows up $m-s$ times in the final expression with a negative sign, stemming from the negative part, but referring to the corresponding expression for $s+1$.[11] The condition that in the end only the original terms remain, i.e. $\Delta\bar{P}_{T,T+1} = \bar{P}(x_1(T+1), ..., x_i(T+1), ..., x_m(T+1)) - \bar{P}(x_1(T), ..., x_i(T), ..., x_m(T))$, requires coefficients unequal 1 for the various terms. In the simplest case, the coefficients of the positive terms can be chosen to be $\frac{1}{s}$ and for the negative ones $\frac{1}{m-s}$, for $s \neq 0$ and $s \neq m$, and $\frac{1}{m-s} = \frac{1}{m}$ for $s = 0$ while the positive part is absent, and $\frac{1}{s} = \frac{1}{m}$ for $s = m$, where the negative part is absent. This decomposition is symmetric and residual-free.

A more general choice of the coefficients is then $\gamma(m,s)\frac{1}{s}$ and $\gamma(m,s)\frac{1}{m-s}$ with $\gamma(m,0) = 1 = \gamma(m,m)$. Choosing $\gamma(m,s) = \frac{s!(m-s)!}{m!}$ then gives the Shapley-value coefficients (see e.g. Baye (2005), taking $s+1$ instead of $s$ for the negative terms) and the decomposition coincides with the Shapley-value based poverty decomposition as introduced in Shorrocks (1999). For two variables, this is equivalent to equation (8). The specific choice of coefficients for the Shapley-value is motivated by symmetry arguments and the Shapley-value has some distinct axiomatic background (symmetry, additivity, no distinguished variable), but I will show in the next subsection that the Shapley-value is not optimal in the light of decomposition as integral

[11] $\pi_{s-1,m-s}$ gives $s$ variables at $T+1$ in the positive parts of $\Delta P^{x_i}_{T,T+1}(\pi_{s-1,m-s})$ and $s-1$ at $T+1$ in the negative ones. Correspondingly, the term with $s+1$ instead of $s$, i.e. $\pi_{s,m-s-1}$, gives $s$ variables at $T+1$ in the negative parts that combine with the corresponding terms from the positive part with $s$.

approximation, and that these axioms cannot be employed as a motivation for the method's optimality.

There are other approaches aiming at improving poverty decomposition. For a recent review of methods, see e.g. Heshmati (2004). Here, I shortly present some illustrative examples. Thorbecke and Jung (2006) base their decomposition on the linkages captured in the Social Accounting Matrix (see also references therein). Second, Fournier (2001) discusses an approach explicitly taking into account changes in the different underlying variables and their correlations separately. This is, in fact, similar to taking some terms of the Shapley-value approach into account and explaining part of the remaining residual by building counterfactuals based on the rank-correlation structure. Usually, some residual remains. Thirdly, there are regression-based approaches to decomposition (see e.g. Juhn et al. (1993), Borooah (2005), Wan and Zhou (2005), Dercon (2006) or Wan and Zhang (2008) and references therein). The regressions, however, refer to the definition, choice, or identification of the variables the decomposition is based on or the construction of the counterfactual case, while the decomposition itself (i.e. the combination of the terms where only one variable changes) is again made according to the common approaches as described in this or the previous subsection. Similarly, Di Nardo et al. (1996) discuss a kernel estimation approach to construct counterfactuals needed for the decomposition into changes attributable to single variables, while the decomposition itself is again a variant of the approaches discussed above.

## 3.3 Poverty Decomposition and Integral Approximation

In this subsection, I discuss the poverty decomposition approaches introduced above in the light of general decomposition as integral approximation as presented in section 2. This establishes a common basis for and a new understanding of poverty decomposition methods.

### 3.3.1 Most Common Approaches and the Shapley-Value

Approximating the terms in equation (5) by their values at the upper boundary leads to expressions such as $J = \frac{\partial P}{\partial \mu}\frac{\partial \mu}{\partial t}|_{T+1}\Delta T$, and approximating the derivatives by the slope of the straight line joining the end-points as discussed in footnote 7 gives[12]

$$J = \frac{P(\bar{\mu}(T+1), \bar{L}(T+1)) - P(\bar{\mu}(T), \bar{L}(T+1))}{\bar{\mu}(T+1) - \bar{\mu}(T)} \frac{\bar{\mu}(T+1) - \bar{\mu}(T)}{\Delta T} \Delta T, \tag{10}$$

which is the Laspeyres index. The corresponding expression can be calculated for the variable $\bar{L}$ and both can also be evaluated at time $T$, thus giving the Paasche index. The combination of the Laspeyres for both $\bar{\mu}$ and $\bar{L}$ gives the Datt-Ravallion decomposition equation (6), and the combination of Laspeyres for $\bar{\mu}$ and Paasche for $\bar{L}$ gives the Jain-Tendulkar formula (7). Taking the average of the Laspeyres and Paasche indices gives the Marshall-Edgeworth index. This, finally, is the same as the Shapley-value decomposition for two variables, equation (8).

[12] Here, the derivatives of $P$, $\frac{\partial P}{\partial \bar{\mu}}$ and $\frac{\partial P}{\partial \bar{L}}$, are approximated by the boundary values of $P$, although these derivatives are known, as the functional form of $P$ is known, cf. page 8.

So far, I have shown how the basic poverty decomposition methods can be seen as special cases of integral approximation. This is however not true any longer for the generalised formulae used in the literature and presented above, i.e. for the Shapley-value with more than two variables. One criticism is that in the light of the equivalence of the Shapley-value decomposition and the decomposition method introduced in Sun (1998) (Ang et al. 2003), the various terms in the Shapley-value can be understood as an assignment of the residual to the various effects based on some symmetry arguments but without further basis in the properties of the underlying functions or integral approximations. Thus, all variables are treated equally, irrespective of their properties. I illustrate this for three variables and a total which is their multiplication:

$$\begin{aligned}\Delta P = P(T) - P(0) &= x_1(T)x_2(T)x_3(T) - x_1(0)x_2(0)x_3(0) \\ &= \Delta P_1 + \Delta P_2 + \Delta P_3, \end{aligned} \tag{11}$$

where $\Delta P_i$ is the contribution of the variable $x_i$ to the decomposition of $P$. Replacing $x_i(T)$ with $x_i(0) + \Delta x_i$, seeing $\Delta x_i$ as the incremental change in $x_i$ from period 0 to $T$, and symmetrically rearranging terms, we have

$$
\begin{aligned}
\Delta P &= (x_1(0)+\Delta x_1)(x_2(0)+\Delta x_2)(x_3(0)+\Delta x_3) - x_1(0)x_2(0)x_3(0) = \\
&= \Delta x_1 x_2(0)x_3(0) + \Delta x_2 x_1(0)x_3(0) + \Delta x_3 x_1(0)x_2(0) + \qquad (12)\\
&\quad +\Delta x_1\Delta x_2 x_3(0) + \Delta x_1\Delta x_3 x_2(0) + \Delta x_2\Delta x_3 x_1(0) + \Delta x_1\Delta x_2\Delta x_3 = \\
&= \Delta x_1 x_2(0)x_3(0) + \frac{1}{2}[\Delta x_1\Delta x_2 x_3(0) + \Delta x_1\Delta x_3 x_2(0)] + \frac{1}{3}\Delta x_1\Delta x_2\Delta x_3 + \\
&\quad +\Delta x_2 x_1(0)x_3(0) + \frac{1}{2}[\Delta x_1\Delta x_2 x_3(0) + \Delta x_2\Delta x_3 x_1(0)] + \frac{1}{3}\Delta x_1\Delta x_2\Delta x_3 + \\
&\quad +\Delta x_3 x_1(0)x_2(0) + \frac{1}{2}[\Delta x_1\Delta x_3 x_2(0) + \Delta x_2\Delta x_3 x_1(0)] + \frac{1}{3}\Delta x_1\Delta x_2\Delta x_3.
\end{aligned}
$$

The three last lines are $\Delta P_1$, $\Delta P_2$ and $\Delta P_3$, respectively, and equal the contributions of the three variables as identified in Sun (1998). As shown in Ang et al. (2003), they are equal to the Shapley-value decomposition, as can also be seen by further rearranging terms and comparing to the formulae for the Shapley-value given above. As already indicated, the logic behind this formula is to equally assign all the difference-terms involving $\Delta x_i$'s to the contributions of the variables $x_i$, i.e. a term involving $s$ $\Delta$-factors is divided by $s$. An illustration for this simple example are the volumes of two cuboids with edges $x_i(0)$ and $x_i(T) = x_i(0) + \Delta x_i$ $(i = 1, 2, 3)$, respectively, and how to assign the difference in volume between the two to each of the differences in the single edges (cf. figure 1).

We are not primarily interested in $\Delta x_i$ itself, but rather in $\Delta x_i = x_i(t + \Delta t) - x_i(t)$ *as a function of* t or, respectively, $\Delta t$. In general, $\Delta x_i$ will be different functions of $\Delta t$ for different $i$, and imposing equal treatment of all $\Delta x_i$ is thus adequate in certain special cases only. I illustrate this criticism of the Shapley-value with a simulation based on some concrete choice of the variables $x_i$ as functions of $t$: choose $x_1 = t, x_2 = t^2, x_3 = \frac{t}{4}$. Inserting this in equation (2), where again $P = x_1x_2x_3$, and solving the integrals gives $\Delta P = \Delta P_1 + \Delta P_2 + \Delta P_3 = \frac{T^4}{4}$ and the following (exact) decomposition

$$\Delta P_1 = \int_0^T \frac{\partial x_1}{\partial t} x_2 x_3 \mathrm{d}t = \int_0^T \frac{t^3}{4} \mathrm{d}t = \frac{T^4}{16},\ \Delta P_2 = \frac{T^4}{8}, \Delta P_3 = \frac{T^4}{16}. \tag{13}$$

Using the Shapley-value equation (12), the result is different (but also exact), which shows that the Shapley-value does not necessarily lead to the correct decomposition[13]:

$$\Delta P_1 = \Delta P_2 = \Delta P_3 = \frac{T^4}{12}. \tag{14}$$

For further illustration, I also state the condition for the Shapley-value for three variables to be exact as an integral approximation. It is, for the contribution of the first variable, the requirement that

$$\begin{aligned}
&\int_0^T \frac{\partial x_1(t)}{\partial t} x_2(t) x_3(t) \mathrm{d}t \stackrel{!}{=} \\
\stackrel{!}{=}\ & \big[x_1(T) - x_1(0)\big] x_2(0) x_3(0) + \\
&+\frac{1}{2}\big[x_1(T) - x_1(0)\big]\big[x_2(T) - x_2(0)\big] x_3(0) + \\
&+\frac{1}{2}\big[x_1(T) - x_1(0)\big]\big[x_3(T) - x_3(0)\big] x_2(0) + \\
&+\frac{1}{3}\big[x_1(T) - x_1(0)\big]\big[x_2(T) - x_2(0)\big]\big[x_3(T) - x_3(0)\big].
\end{aligned} \tag{15}$$

Comparing this to integral approximation as discussed above shows that the Shapley-value contains too many terms mixing values referring to the two different boundaries. In correct integral approximation, for each additive

[13] Most terms are equal zero in this simple example, as $x_i(0) = 0$ for $i = 1, 2, 3$, but this special property is not crucial for the general argument.

contribution, such mixture only occurs via the derivative-term, i.e. for one variable only, while all the others are evaluated *either* at the upper *or* lower boundary only (cf. page 9).

### 3.3.2 Axiomatic Decomposition

Here, I link the poverty decomposition method based on integral approximation as described above to some axiomatic approaches in the literature. A recent example is Kakwani (2000), who sets up a system of 5 simple rather intuitive axioms any poverty decomposition should fulfill[14], discusses and criticises existing decomposition methods in the light of these axioms and proposes a new method that fulfills all 5 axioms. His discussion is framed in a two-variable setting and the method he finally recommends is just the Shapley-value for two variables.[15] As can be seen from direct calculation, due to the properties of integration, the basic formula for the decomposition based on integration approximation, equation (3), fulfills the 5 axioms set up by Kakwani (2000).

Other axiomatic systems are presented in Shorrocks (1982), Paul (2004) and Tsui (1996), for example. The axiomatisation in Tsui, however, mainly refers to the poverty measure itself and less to its decomposition, which is

[14]As usual, the change of a poverty measure $P$ is decomposed into a growth and an inequality component: $\Delta P_{ij} = G_{ij} + I_{ij}$ for periods $i$ and $j$. The axioms are 1) If $I_{ij} = 0$ then $\Delta P_{ij} = G_{ij}$ and if $G_{ij} = 0$ then $\Delta P_{ij} = I_{ij}$; 2) if $G_{ij} \leq 0$ and $I_{ij} \leq 0$ then $\Delta P_{ij} \leq 0$ and if $G_{ij} \geq 0$ and $I_{ij} \geq 0$ then $\Delta P_{ij} \geq 0$; 3) $G_{ij} = -G_{ji}$ and $I_{ij} = -I_{ji}$; 4) $G_{ij} = G_{ik} + G_{kj}$; 5) $I_{ij} = I_{ik} + I_{kj}$ for all periods $i, j, k$;

[15]He does not mention this, though - but the Shapley-value decomposition was also introduced after this paper was originally written in 1997.

basically the same as the one finally derived in Kakwani (2000).[16]

The axioms in Shorrocks (1982) and Paul (2004) primarily refer to the static decomposition of (income) inequality into different income types.[17] They are thus not of direct relevance for the dynamic decomposition considered here. Nevertheless, I point out that in these approaches, some symmetry properties are important, which also play a role in Shorrocks (1999), but are not reflected in the basic formula (3) of decomposition based on integral approximation. This mismatch of symmetries in the Shapley-value and the approach based on integral approximation has already been illustrated above.

Two conclusions from this discussion on axiomatic approaches may be drawn. First, equation (3), the basic formula of decomposition based on integral approximation fulfills some set of axioms (e.g. Kakwani 2000) and can thus be seen as one realization of an axiomatic approach. Second, when it comes to concrete approximation, though, axioms should not be given too much weight to. This is because of incomplete information on the development of the variables decomposition is based upon. The concrete implementation of decomposition necessitates taking the limited information on the variables involved into account, which is reflected in the necessity to undertake approximations. During this step, from exact formulation via integrals towards concrete calculations via approximation of integrals, some axiomatic

[16] I emphasize that axiomatic foundations of poverty measures and axiomatic foundations of decomposition methods must not be confused. Here, I am concerned with the latter only.

[17] Paul (2004) critizises Shorrocks (1982) for the lack of motivation for some of his axioms. He retains three of Shorrocks' axioms and adds two new ones. Fournier (2001) also critizises Shorrocks' axioms as being too restrictive and, being static, as not being of primary interest — although static decomposition is applied frequently.

properties may be lost. Concrete decomposition thus need not to exactly fulfill the axioms. Although axioms may inform the general structure of some decomposition approach, due to the presence of incomplete information, they can be violated in application.[18]

# 4 Conclusions

A wide range of methods for poverty or general inequality measure decomposition is currently being applied. None of these methods, however, has a sound basis, as none refers to integral approximation, which is the ultimate starting point of any dynamic decomposition analysis. The Shapley-value, for example, assigns the residual term in an inadequate manner to the different drivers behind changes in poverty. This does not mean that results based on the Shapley-value are necessarily wrong - but it is difficult to assess when it is adequate and how large potential errors may be. To assess the adequacy of the methods most often applied in poverty decomposition, such as the Shapley-value, comparison with methods more directly related to integral approximation is necessary.

Muller (2007) provides some preliminary analysis of these issues in the context of energy and pollutant decomposition, where similar problems are encountered. There, the Logarithmic Mean Divisia Index LMDI (Ang 2004), is identified as a method that performs reasonably well also in relation to in-

[18]However, being an approximation to formulae that fulfill axioms, a concrete decomposition thus may be expected to also "approximately" fulfill the axioms, i.e. to not show "too large deviations" from fulfillment. Thus, some non-zero residual is not problematic, but if the residual dominates effects, the decomposition is not of much use.

tegral approximation, although also lacking a sound theoretical basis related to this. It is thus promising to also use the LMDI in poverty decomposition. A more in-depth assessment is necessary, though.

Strategies to proceed with such - and with the assessment of the performance of any decomposition method in relation to integral approximation - are to investigate the performance of the method in simulations, where the exact results are known (cf. the example on page 19), and to identify classes of functions, for which the proposed method is exact or a good approximations if compared to the exact solution. This needs to be combined with all knowledge available on the functions that are being decomposed. The functional form of the poverty measure is, for example, known. The functional form of Its derivatives thus need not be approximated. Furthermore, for short time intervals, for example, some linearity assumptions may be reasonable (see Bresson (2008) for such a strategy applied to the Shapley value in combination with integral approximation). There may also be some information on seasonal or other patterns for the periods between the points where the functions are known. Often, the functions involved will also be based on one-period-back average or aggregate values (e.g. income at time $T$ is the aggregate annual income from $T-1$ to $T$), a property that may be exploited.

Further research is also needed on how the decomposition based on integral approximation relates to certain types of static decomposition, e.g. regarding different types of income sources. Given some time-development, accounting for group structure and for different types of income is possible without problems[19], but within a truly static setting, some reformulation of

[19] For different income types, one could consider an inequality measure depending on

this approach may be necessary. On the other hand, one may argue that a static decomposition without time development is of secondary interest only.

Finally, reliance on axiomatic approaches is no solution to identify optimal methods. In the light of integral approximation, desirable properties only need to be fulfilled approximately. The prime example for this is the desirability of a zero residual, i.e. of a complete decomposition, which does not need to hold for an approach based on approximations. It is only natural to encounter some errors when approximating - which simply lies in the nature of an approximation in comparison to an exact solution. A zero residual thus bears the danger of having been forced to be zero by just randomly or without strong basis apportioning it to the different parts of a decomposition. A decomposition with zero residual thus needs not be superior to one with a non-zero residual.

# References

Ang, B.W., 1995. Decomposition methodology in industrial energy demand analysis. Energy 20(11), 1081-1095.

Ang, B.W., 2004. Decomposition analysis for policymaking in energy: which is the preferred method? Energy Policy 32, 1131-1139.

Ang, B.W., Liu, F.L., Chew, E.P., 2003. Perfect Decomposition Techniques in Energy and Environmental Analysis. Energy Policy 31, 1561-1566.

the different sets of parameters (mean, variance, etc. ) characterizing the distribution of each income type.

Balk, B., 2005. Divisia price and quantity indices: 80 years after. Statistica Neerlandica 59(2), 119-158.

Barro, R., Sala-I-Martin, X., 2003. Economic Growth, 2nd edition, MIT Press.

Baye, F., 2004. Growth and Redistribution Effects of Poverty Changes in Cameroon: A Shapley Decomposition Analysis. In: African Development and Poverty Reduction: The Macro-Micro Linkage, South Africa. See also Baye, F., 2006, Growth, Redistribution and Poverty Changes in Cameroon: A Shapley Decomposition Analysis, Journal of African Economies 15(4): 543-570.

Baye, F., 2005. Structure of Sectoral Decomposition of Aggregate Poverty Changes in Cameroon. In: Proceedings of the International Conference on Shared Growth in Africa, Accra, Ghana.

Bhanumurthy, N., Mitra, A., 2003. Declining Poverty in India: A Decomposition Analysis. Working Paper.

Borooah, V., 2005. Caste, Inequality, and Poverty in India. Review of Development Economics 9(3), 399-414.

Bresson, F., 2008. The Estimation of the Growth and Redistribution

Components of Changes in Poverty: A Reassessment. Presentation at the UNU-WIDER Conference on Frontiers of Poverty Analysis, Helsinki 2008.

Bruvoll, A., Larsen, B.M., 2004. Greenhouse gas emissions in Norway: do carbon taxes work? Energy Policy 32, 493-505.

Chakravarty, S., Deutsch, J., Silber, J., 2008. On the Watts Multidimensional Poverty Index and its Decomposition. World Development 36(6): 1076-1077.

Datt, G., Ravallion, M., 1992. Growth and Redistribution Components of Changes in Poverty Measures: A Decomposition with Application to Brazil and India in the 1980s. Journal of Development Economics 38(2), 275-295. Working Paper version from 1991, Living standards measurement study working paper no. 83. The World Bank, Washington, DC. http://www-wds.worldbank.org/external/default/WDSContentServer/ WDSP/ IB/ 1999/ 09/ 17/000178830_ 98101902173994/ Rendered/ PDF/ multi_ page. pdf (10.11.2008).

Dercon, S., 2006. Economic Reform, Growth and the Poor: Evidence from Rural Ethiopia. Journal of Development Economics 81(1), 1-24.

Di Nardo J., Fortin, N., Lemieux, T., 1996. Labour market institutions and the distribution of wages 1973-1992 : A semiparametric approach. Econometrica 64(5), 1001-1064.

Diewert, W., Nakamura, A. (eds), 1993. Essays in Index Number Theory, Volume I, Elsevier Science Publishers. http://www.econ.ubc.ca/ diewert/ inbook.htm (23.10.2008)

Duclos, J.-Y., Araar, A., Giles, J., 2006. Chronic and Transient Poverty: Measurement and Estimation, with Evidence from China, IZA Discussion Papers No 2078, Institute for the Study of Labor (IZA). http:// econpapers. repec. org/ paper/izaizadps/dp2078.htm (22.10.2008).

Foster, J., Greer, J., Thorbecke, E., 1984. A Class of Decomposable Poverty Measures. Econometrica 52(3), 761-766.

Fournier, M., 2001. Inequality Decomposition by Factor Component: A "Rank-Correlation" Approach Illustrated on the Taiwanese Case. Recherches Économiques de Louvain - Louvain Economic Review 67(4), 381-403.

Grootaert, C., 1995. Structural Change and Poverty in Africa: A Decomposition Analysis for Côte d'Ivoire. Journal of Development Economics 47(2), 375-401.

Heshmati , A., 2004. A Review of Decomposition of Income Inequality. IZA Discussion Paper No. 1221. http:// ssrn.com/ abstract= 571703 (10.11.2008).

Jain, L., Tendulkar, S., 1990. The Role of Growth and Distribution in the Observed Change in Head-Count Ratio-Measure of Poverty: A Decomposition Exercise for India. Indian Economic Review XXV(2), 165-205.

Juhn C., Murphy, K., Pierce, B., 1993. Wage Inequality and the Rise in Returns to Skill. Journal of Political Economy 101(3), 410-442.

Kakwani, N., 2000. On Measuring Growth and Inequality Components of Poverty with Application to Thailand. Journal of Quantitative Economics 16(1), 67-79. Working Paper Version (1997): Discussion Paper 97/16, School of Economics, The University of New South Wales, ISBN 0 7334 1576 8.

Kolenikov, S., Shorrocks, A., 2005. A Decomposition Analysis of Regional Poverty in Russia. Review of Development Economics 9(19), 25-46.

Kraay, A., 2006. When is Growth Pro-Poor? Evidence from a Panel of Countries. Journal of Development Economics 80(1), 198-227.

Mazumdar, D., Son, H., 2001. Vulnerable groups and the Labor Market in Thailand. Working Paper.

Morduch, J., Sicular, T., 2002. Rethinking Inequality Decomposition, with Evidence from Rural China. The Economic Journal 112: 93-106.

Muller, A., 2007. Putting decomposition of energy use and pollution on a firm footing - clarifications on the residual, zero and negative values and strategies to assess the performance of decomposition methods. Working Papers in Economics Nr 213 and 215 (combined, revised version from 10.8.2007), Dept. of Economics, Göteborg University. http://hdl.handle.net/2077/2700 (10.11.2008).

Paul, S., 2004. Income Sources Effects on Inequality. Journal of Development Economics 73(1), 435-451.

Shorrocks, A., 1982. Inequality Decomposition by Factor Components. Econometrica 50(1), 193-212.

Shorrocks, A., 1999. Decomposition Procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value. Working Paper.

Shorrocks, A., 2008. Poverty Decompositions. Keynote lecture at the UNU-WIDER Conference on Frontiers of Poverty Analysis, Helsinki 2008.

Son, H., 2003. A New Poverty Decomposition. Journal of Economic Inequality 1, 181-187.

Sun, J.W., 1998. Changes in energy consumption and energy intensity: a complete decomposition model. Energy Economics 20, 85-100.

Thorbecke, E., Jung, H.-S., 1996. A Multiplier Decomposition Method to Analyze Poverty Alleviation. Journal of Development Economics 48(2), 279-300.

Trivedi, P.K., 1981. Some discrete approximations to Divisia integral indices. International Economic Review 22(1), 71-77.

Tsui, K., 1996. Growth-equity decomposition of a change in poverty: an axiomatic approach. Economics Letters 50, 417-423.

Wan, G., Zhou, Z., 2005. Income Inequality in Rural China: Regression-based Decomposition Using Household Data. Review of Development Economics 9(1), 107-120.

Wan, G., Zhang, Y. 2008. Explaining the Poverty Difference between Inland and Coastal China: A Regression-Based Decomposition Approach. Review of Development Economics 12(2), 455-67.

Yun, M.-S. 2004. Decomposing Differences in the First Moment. Economics Letters 82, 275-280.

# Figure Captions

Figure 1: Illustration of the basic task in decomposition

Figure 2: Step-functions for integral approximation. The grey rectangle is the Paasche method, this plus the dashed rectangle give Laspeyres, the grey

rectangle and the dotted triangle Marshall-Edgeworth