UNITED NATIONS
UNIVERSITY
UNU-WIDER

WIDER Working Paper 2021/142

The recentered influence function and unidimensional poverty measurement

Carlos Gradín*

September 2021

United Nations University World Institute for Development Economics Research
wider.unu.edu

Abstract: I discuss the applicability of the recentered influence function (RIF) to the analysis of poverty differentials between distributions (regression-based decomposition into composition and income structure effects). I show that the predominant approach in the empirical literature estimates the relationship between individual poverty functions of additive measures, particularly the head-count ratio, and household attributes. Given that the recentered influence function of these measures is also their poverty function, this approach is simply a specific case of the one-stage recentered influence function decomposition, using non-linear probability models. However, the use of recentered influence function provides a more general approach that better accounts for individual contributions to poverty for non-additive poverty measures (such as that of Sen and its extensions) as well. At the same time, the use of reweighting in a first stage allows to avoid imposing any functional form on the relationship between poverty and characteristics at the aggregate level.

Key words: poverty, recentered influence function (RIF), regression-based decomposition

JEL classification: C46, I32

Acknowledgements: This study has been prepared within the UNU-WIDER project World Income Inequality Database (WIID), as part of the *Handbook of Research on Measuring Poverty and Deprivation*, edited by Jacques Silber for Edward Elgar Publishing. The Wellbeing, Income Dynamics and Segregation project (Ministerio de Ciencia e Innovación, grant PID2019-104619RB-C41) is also acknowledged.

---

*UNU-WIDER, gradin@wider.unu.edu

This study has been prepared within the UNU-WIDER project World Income Inequality Database (WIID).

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9267-082-5

https://doi.org/10.35188/UNU-WIDER/2021/082-5

Typescript prepared by Lorraine Telfer-Taivainen.

United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

In recent decades, the empirical economic literature has witnessed a growing interest in the use of regression-based decomposition analyses to understand the drivers of distributional changes over time, as well as the distributive differences between population groups, countries, or periods. The main purpose of these analyses is to decompose a differential in a statistic of interest (for example, the mean) between two distributions (e.g., male and female, White and non-White, two years) in composition and structural effects (aggregate decomposition). The first effect reflects the part of the differential that is due to the fact that both populations diverge in terms of their composition by relevant characteristics, while the second one reflects the differential that remains conditional on both populations having the same distribution of those characteristics. These two aggregate effects can be further decomposed into the contribution of each explanatory factor or set of factors (detailed decomposition), highlighting the relevance of each attribute to explain the differential (through each effect).

The approach became popular in labour economics after the seminal contributions of Blinder (1973) and Oaxaca (1973) to the analysis of wage discrimination using linear regression. The main idea was to estimate the average wage difference between African Americans and Whites or between men and women that remains after controlling for differences in the distribution of relevant attributes between the two groups of workers (for example, in level of education or experience). The wage gap associated with different returns to worker attributes (that is, not explained by characteristics) can be interpreted as wage discrimination if the main sources of productivity-based differences have been adequately controlled. However, the presence of endogeneity issues or unobservable key attributes (e.g., quality of education, non-cognitive skills, etc.) can pose a challenge in that interpretation. As Oaxaca (1973: 708) points out, the method has deep roots: 'This methodological technique is found in other studies as well and may take the form of regression analysis or standardization analysis'. Indeed, there was, for example, a long tradition of comparing standardized rates between groups in social research. Following this line of the literature, Kitagawa (1955) discussed a non-parametric method to study the differences between the total rates of two groups considering a small number of categorical explanatory factors, a refined and revised version of a technique that had been used since 1948 at the University of Chicago.[1]

The subsequent literature that followed these initial approaches has tried to refine these techniques to address issues that affect the interpretation of the decomposition, such as self-selection in the sampled group, endogeneity or unobserved heterogeneity, among others. This regression-based decomposition approach was later extended to study the differences between two distributions in specific quantiles, such as p25, p50 (the median), or p75, for example, or in aggregate inequality measures such as the variance or the Gini index, or to the entire distribution (for example, density or cumulative distribution).

Fortin et al. (2011) provide a detailed technical discussion of the main different approaches and their econometric properties, highlighting the advantages of combining the estimation of a counterfactual distribution using propensity score reweighting (DiNardo et al. 1996) to obtain the aggregate decomposition, and the approach based on the statistical concept known as recentered influence function (RIF) (Firpo et al. 2007, 2009, 2018) to obtain the detailed decomposition under

[1] A rate is the average of a dummy variable (where 1 indicates group membership, e.g., male or white). A non-parametric method is similar to an OLS with categorical variables including all possible interactions.

simplifying assumptions (linear relationship between RIF and explanatory factors). This approach can be seen as a generalization of the conventional Blinder-Oaxaca framework, which would be a particular case when the statistic is the mean. Although this literature falls within the causal inference analysis in non-randomized treatments using observational data, the conditions for causal identification are rarely met, and therefore these studies generally refer to statistical associations. But even if they are primarily interpreted as observational studies, they have proven particularly useful in gaining a better understanding of distributional phenomena across many dimensions between groups and over time.

The purpose of this paper is to provide an overview of the implications of adapting the RIF approach (with or without reweighting) to the measurement of poverty, which has great potential even if it has rarely been used as such in the literature. As discussed here, to a large extent, the regression-based decompositions of poverty that can be found in the literature can be seen as specific cases of this approach.

Next, I first describe the regression-based decomposition used in the context of poverty. After that, I will analyse the RIF approach applied to poverty measures and explore the relationship between the RIF and individual poverty functions. The final section concludes.

2 Regression-based decomposition of poverty measures

The case of poverty, especially the more popular measure known as the poverty rate or headcount ratio, has not been an exception to the trend of using regression-based decomposition methods in its analyses. Topics such as changes in poverty over time or the poverty gap between two ethnic or racial groups were commonly addressed using this type of approach.[2] Since instead of using individual earnings, these analyses are based on household income or consumption, the possible connection with interpretations in terms of discrimination tends to be weaker. In general, if a population group is discriminated, this will normally have an effect through both effects if its members have limited access to the accumulation of human or physical capital, to jobs with better employment opportunities, family planning, etc., as well as if they receive lower remuneration to their endowments. In this context, the identification of different sources for the differentials gained more relevance to help understand the role of differences in composition by education, type of household, employment status, occupation, or other sociodemographic factors. Alternatively, these differences are attributed rather to different levels of conditional poverty between groups with similar characteristics.

As mentioned above, a direct application of conventional approaches to differences in means can produce a decomposition of two poverty rates using a non-parametric approach or a linear probability model (LPM) if the variable of interest is a dummy variable with 1 reflecting that someone is poor and 0 is not poor. However, the empirical literature has been more inclined to use binary choice models (probit/logit) to estimate the relationship between poverty status and explanatory factors.[3]

Therefore, empirical poverty analyses using regression-based decompositions became more popular as the approach was extended to deal with non-linear probability models. Specially to obtain the detailed effects, which are simple in the linear case but there is no unique or unanimously

[2] E.g., see review of earliest studies in Gradín (2009).

[3] There is still controversy about the use of LPM for binary dependent variables (e.g., Breen et al. 2018).

accepted solution in the case of the non-linear models. Among these approaches, the linearized method proposed by Even and Macpherson (1990, 1993) stands out to obtain an aggregate decomposition and the detailed decomposition of the composition effect. This was later formalized and generalized by Yun (2004) to include the detailed structure effect as well.[4] I will focus on this approach due to its popularity and direct connection to RIF, as explained later. The most common practice was to estimate models in a reduced form, rather than structural models, given the complexity of modelling all the economic and sociodemographic factors that affect household poverty.

3 Poverty measures and poverty functions

Amartya Sen has well established that the study of poverty requires two stages, namely, identification of the poor, and the quantification of aggregate poverty (Sen 1976). The identification of the poor in a unidimensional setting is usually done by considering poor people with income or consumption below a specific poverty line z, which can generally be obtained as an absolute measure of needs (i.e., the cost of purchasing a basic bundle of food and non-food items) or a percentage (e.g., 50 or 60) of a baseline income (e.g., the median). In the second step, for a given poverty line, overall poverty is aggregated using poverty indices, that can generally be written as a functional of the cumulative distribution of income y (e.g., household net income per capita): $P(F_y; z)$. In general, we can think of y here as the censored income distribution where the incomes of non-poor people are assigned a value z. This is the result of the generally accepted 'focus axiom', implying that a change in the income of the non-poor does not affect overall poverty, provided they do not fall below the poverty line. For simplicity, in what follows, the notation of poverty measures will ignore z, which is taken as given.

In the context of regression-based analyses, it is also important to note that most common poverty measures can also be represented as the average across the population of individual poverty functions: $P(F_y) = E\left(p(F_y)\right)$, with $p(F_y)$ being the poverty function, that takes positive values for the poor and a value of zero for the non-poor.

When the individual poverty function for any individual income y depends only on how far the income is from the poverty line (and do not depend on other people's incomes), overall poverty measures are additively decomposable: $P(F_y) = E(p(y))$. This is the case of the most popular family of indices used in the empirical literature, proposed by Foster, Green and Thorbecke (1984) (FGT), where the poverty function is a transformation of the normalized distance between each income and the poverty line: $p_\alpha(y) = \left(\frac{z-y}{z}\right)^\alpha$. The family includes as particular cases well-known indices such as the head-count ratio, the average normalized poverty gap, and poverty severity (average squared normalized poverty gap) for $\alpha = 0,1,2$. There are other additive indices, such as the first known distributive-sensitive poverty measure, proposed by Watts (1969), where the poverty function can be interpreted as the loss of welfare due to poverty: $p(y) = lnz - lny$.

In other cases, however, individual poverty functions may also depend on other people's incomes. This is the case of the two measures proposed by Sen (1976) and the various extensions that can

[4] There are other strategies to address non-linearity: e.g., sequential decomposition (Gomulka and Stern 1990; Fairlie 1999, 2005) or evaluating differences in characteristics at marginal effects (Schwiebert 2015).

be found in the literature,[5] where individual poverty functions weight each normalized poverty gap by a measure of people's rank within the poor or within the entire population. In cases like these, poverty measures do not verify additive decomposability, or even a weaker property of subgroup consistency (according to which overall poverty is increasing in poverty in any group or individual, ceteris paribus).[6] Although there is a rationale for not verifying this property, some people consider that this feature makes these indices less attractive for empirical analyses.

4 Aggregate decomposition of a difference in poverty measures

The aggregate decomposition of poverty between two distributions ($t = 0, 1$), that could be two populations or two years, aims at identifying in such differential a composition effect driven by differences in the distribution of key households' characteristics (an $N_t xK$ matrix X_t) observed with a common support in both distributions, and an income structure effect that reflects how the conditional distribution of income of people with those given characteristics, $F_{y_t} \equiv F(y_t|X_t)$, differs between both distributions (poverty conditional on characteristics).[7]

Let $P(F(y_s|X_t))$ be the poverty measure when people in a distribution t obtain incomes under the income structure prevailing in distribution s. Then $P(F(y_0|X_0))$ and $P(F(y_1|X_1))$ are the corresponding observed poverty measures in each comparison distribution. In this context, it will be useful to also consider $P(F(y_0|X_1))$ as the poverty level that would have prevailed in the counterfactual or hypothetical situation in which people in distribution 1 obtained their income under the income structure of 0.[8] By just adding and subtracting this counterfactual measure to the overall difference Δ_O^P, we get the aggregate decomposition into the income structure effect Δ_S^P (differential in conditional poverty evaluated with the distribution of characteristics in 1) and the composition effect Δ_X^P (poverty differential induced by a different composition of the population by characteristics, evaluated at conditional poverty levels observed in 0):

$$P(F(y_1|X_1)) - P(F(y_0|X_0)) = [P(F(y_1|X_1)) - P(F(y_0|X_1))] + [P(F(y_0|X_1)) - P(F(y_0|X_0))]$$

$$\Delta_o^P \quad = \quad \Delta_S^P \quad + \quad \Delta_X^P$$

To undertake this aggregate decomposition, it is necessary to estimate the counterfactual $P(F(y_0|X_1))$. Most of the literature on poverty analysis has done this by assuming a relationship between individual poverty functions of the FGT family (mainly the head-count ratio) and a linear combination of the row vector of individual household characteristics x, $p_\alpha(y) = g(x\beta)$, allowing to estimate the counterfactual measure as the average predicted value for one group (t)

---

[5] E.g., see Zheng (1997) or Foster (2006: 41-65).

[6] Additive decomposability is more demanding that subgroup consistency, since it imposes proportionality between the increase in group and overall poverty levels. Most measures in Zheng (1997) that are subgroup consistent are also additively decomposable.

[7] These effects receive different names: characteristics or explained (composition); coefficients or unexplained (structure).

[8] There is an index number problem. The roles of both distributions in the counterfactual can be reversed. The decomposition can also be averaged over the two possible alternatives (Shapley decomposition; e.g., Shorrocks 2013) or estimated using the pooled sample, among other options (see Fortin et al. 2011; Jann 2008).

using the estimated coefficients for the other group (s): $P(F(y_s|X_t)) = E\left(g(x_t\beta_s)\right) = \bar{g}(x_t\beta_s)$. A similar logic could be applied to the poverty functions of other indices.

The above relationship can be linear, like in the Blinder-Oaxaca approach, leading to the estimation of the LPM in the most popular case of the head-count ratio, for example, or the ordinary least squares (OLS) estimation of any transformation of the normalized poverty gap. However, the relationship may also be non-linear in other cases, like when using logit or probit in the case of the head-count ratio, imposing that the predicted probabilities fall between 0 and 1. Or when using a tobit model in the case of the poverty gap or its square, for example, taking into account that normalized poverty gaps are actually censored at the poverty line, when they take value 0.

Alternatively, one may consider adapting the RIF decomposition approach to this context, as will be explained in more detail below. I first discuss the RIF concept, and then the two possible strategies for implementing the RIF approach: one-stage or two-stage decomposition, depending on whether the aggregate decomposition is first obtained by reweighting.

5 The RIF of poverty measures

The influence function (IF) of a poverty measure $IF(y;P)$ quantifies the impact on poverty of marginally increasing the population mass at a certain income y (i.e., a small 'contamination') and has an expected value of zero. More formally, if F_ε is a mixture distribution assigning a probability $1-\varepsilon$ to the original distribution F and ε to a degenerated distribution with all its population mass at a y, the IF is the directional derivative of P for this mixture distribution when $\varepsilon \to 0$ (see Hampel 1974):

$$IF(y;P) = \frac{\partial}{\partial\varepsilon} P(F_\varepsilon)\,)|_{\varepsilon=0} \text{ ; with } E(IF(y;P) = 0$$

The recentered influence function $RIF(y;P)$ (Firpo et al. 2007) is just obtained after recentering the IF at the value of P:

$$RIF(y;P) = P\left(F_y\right) + IF(y;P)\text{; with } E(RIF(y;P) = P\left(F_y\right)\,)$$

A RIF poverty regression will relate individual RIF values of a specific poverty measure and household characteristics given by X that can be estimated by OLS, so that:

$$P\left(F_y\right) = E\left(RIF(y;P)\right) = \bar{X}\beta$$

$P\left(F_y\right)$ can thus be written as just the weighted sum of the impact of all average household characteristics on poverty. In this context, it is useful to note that the RIF can be interpreted in terms of individual or group contributions to poverty (Gradín 2020).

In the case of additively decomposable measures, it is straightforward to note that the corresponding RIF is just given by the individual poverty function:

$$RIF(y;P) = p(y)$$

This is derived from the fact that individual poverty functions themselves are not affected by the contamination: $P(F_\varepsilon) = (1-\varepsilon)P\left(F_y\right) + \varepsilon p(y)$, and $\frac{\partial}{\partial\varepsilon} P(F_\varepsilon\,)|_{\varepsilon=0} = p(y) - P\left(F_y\right) = IF(y;P)$.

In cases like the Sen measure and extensions in which the poverty function weights the normalized poverty gap by a measure of the individual rank, the RIF should also account for an indirect effect through a change in the rank-based weights of other incomes and will not be equal to the poverty function. Essama-Nssah and Lambert (2012: 135–159) have already obtained the necessary expressions for some of these additively decomposable indices like the FGT or Watts, which are straightforward, but also for less straightforward cases, including the Sen index, as well as for other useful elements of the poverty analysis toolkit like the Three 'I's for Poverty (TIP) ordinates, the poverty elasticity of the headcount ratio, and some pro-poorness measures, that can be used in a similar fashion.

Since the RIF is basically given by the individual poverty functions (plus the indirect effect in the case of rank-based measures), they will generally inherit their properties. For example, the RIF of poverty measures will generally be zero for incomes above the poverty line, will tend to be larger for poorest incomes in the case of monotonic indices, being disproportionally larger for distributive-sensitive measures.

5.1 A one-stage RIF decomposition

The RIF regressions can then be used to decompose Δ_O^P as in the conventional Blinder-Oaxaca decomposition, with individual $RIF(y;P)$ being the dependent variable, obtaining the overall, composition and structure effects as:

$$\Delta_O^P = [\bar{X}_1\beta_1 - \bar{X}_0\beta_0] = \Delta_X^P + \Delta_S^P$$

$$\Delta_X^P = [\bar{X}_1 - \bar{X}_0]\beta_0;\ \Delta_S^P = \bar{X}_1[\beta_1 - \beta_0]$$

Due to linearity, the individual contribution of each average variable $\bar{x}^k$ to each effect can be obtained as:

$$\Delta_{O^k}^P = \bar{x}_1^k\beta_1^k - \bar{x}_0^k\beta_0^k = \Delta_{X^k}^P + \Delta_{S^k}^P$$

$$\Delta_{X^k}^P = \left(\bar{x}_1^k - \bar{x}_0^k\right)\beta_0^k\ ;\ \Delta_{S^k}^P = \bar{x}_1^k\left(\beta_1^k - \beta_0^k\right)$$

This procedure is therefore equivalent to running the Blinder-Oaxaca decomposition of differences in poverty functions of additive measures, using a linear probability model.

As discussed above, the main approach adopted in part of the literature was to estimate non-linear regressions instead (logit/probit or tobit). Using the RIF values as dependent variables, the corresponding effects can be obtained as:[9]

$$\Delta_O^P = \bar{g}(x_1\beta_1) - \bar{g}(x_0\beta_0) = \Delta_X^P + \Delta_S^P$$

$$\Delta_X^P = \bar{g}(x_1\beta_0) - \bar{g}(x_0\beta_0);\ \Delta_S^P = \bar{g}(x_1\beta_1) - \bar{g}(x_1\beta_0)$$

[9] A different approach to deal with non-linearity in the aggregate effects consists of computing poverty using a Probit distribution function of incomes that were predicted using loglinear regressions (Coudouel et al. 2002).

Due to the above linearization, the detailed contributions can be obtained as:

$$\Delta_{O^k}^P = \frac{\bar{x}_1^k \beta_1^k - \bar{x}_0^k \beta_0^k}{\bar{X}_1 \beta_1 - \bar{X}_0 \beta_0} \Delta_O^P = \Delta_{X_k}^P + \Delta_{S_k}^P$$

$$\Delta_{X^k}^P = \frac{(\bar{x}_1^k - \bar{x}_0^k)\beta_0^k}{(\bar{X}_1 - \bar{X}_0)\beta_0} \Delta_X^P; \; \Delta_{S^k}^P = \frac{\bar{x}_1^k(\beta_1^k - \beta_0^k)}{\bar{X}_1(\beta_1 - \beta_0)} \Delta_S^P$$

It is straightforward to see that this is a generalization of the expressions above that are equivalent whenever g is linear because then $\bar{g}(x_t \beta_s) = \bar{x}_t \beta_s$.[10]

A caveat is needed on how to interpret the detailed contributions of the structure effect of these decompositions (Oaxaca and Ransom 1999). For categorical variables (i.e., sets of dummy variables in regressions), the estimated contribution of each category depends on which category was omitted. Also, the contribution of a continuous variable will depend on its scale. The latter is only a problem in the case of variables that do not have a natural or generally accepted scale. For the former, there are some solutions that have been proposed in the literature using a renormalization of the estimated coefficients, but all these are ad hoc (Fortin et al. 2011) and do not change the nature of the problem. This implies that the interpretation of the income structure coefficients should always admit this limitation.

5.2 The two-stage approach (Reweighting + RIF)

In the main complete approach proposed by Firpo et al (2007), the estimation strategy is carried out in two stages. The aggregate decomposition is done in the first stage using a reweighting method approach. This semi-parametric method is based on a propensity score procedure (DiNardo et al. 1996) and provides a consistent estimate of the entire counterfactual distribution under the ignorability assumption (both distributions have the same distribution of unobservables conditional on characteristics), without the need to assume any functional form. Any statistic of interest, including poverty measure P, can be calculated on this counterfactual distribution, obtaining $P(F(y_0|X_1))$, and then both Δ_S^P and Δ_X^P can be easily computed.

If we rewrite the density of a distribution as the integral, over the distribution of individual characteristics, of the product between the density of income conditional on individual characteristics and the marginal distribution of individual characteristics:

$$f(y_t|X_t) = \int_{x \in \Omega_x} f(y_t|x_t) dF(x_t)$$

the counterfactual distribution requires replacing the marginal distribution of X in 0 with the marginal distribution in 1. This can be done by rescaling the sampling weights in 0 by a reweighting factor $\psi(X)$:

$$f(y_0|X_1) = \int_{x \in \Omega_x} f(y_0|x_0) dF(x_1) = \int_{x \in \Omega_x} f(y_0|x_0) dF(x_0) \psi(X)$$

[10] Predicted poverty using probit or tobit is a consistent (but not exact) estimate of the observed value. Thus, the decomposition of the observed gap can be obtained by extrapolation of the decomposition of the predicted gap.

Applying Bayes' rule, this factor can be rewritten as:

$$\psi(X) = \frac{dF(x_1)}{dF(x_0)} = \frac{\Pr(X|t=1)}{\Pr(X|t=0)} = \frac{\Pr(t=1|X)}{\Pr(t=0|X)}\frac{\Pr(t=0)}{\Pr(t=1)}$$

Therefore, to obtain the rescaling factor, we need to estimate the probability that each person is observed in 1 conditional on a set of covariates, $\Pr(t=1|X)$ and its complementary, $\Pr(t=0|X) = 1 - \Pr(t=1|X)$ using a pooled sample of both distributions. In practical terms, this can be done non-parametrically if variables are categorical and the number of categories is not too large for the sample of interest; or using the predicted values of a logit or probit model in which the dependent variable is a dummy indicating membership to distribution 1, and the explanatory variables are the key covariates along all possible interactions. The general idea is to rescale the sampling weights of people in 0 so that their average characteristics are as close as possible to those of 1, while keeping their conditional incomes.

The aggregate decomposition can then be obtained as:

$$\Delta_S^P = X_1\beta_1 - X_R\beta_R;\ \Delta_X^P = X_\mathrm{R}\beta_R - X_0\beta_0,$$

where the subscript R indicates the reweighted distribution. One limitation of this reweighting procedure, however, is that obtaining detailed decompositions, especially of the structure effects, is not straightforward.[11] This is why, in the second stage, the detailed decomposition is obtained using two linear RIF poverty decompositions comparing the counterfactual with each of the two original distributions.

A RIF poverty decomposition of the difference between distribution 1 and the reweighted counterfactual can be used to break the aggregate structural effect into a pure RIF structural effect $\Delta_{S,p}^P$ (the RIF income structure effect in this decomposition) and a RIF reweighing error $\Delta_{S,e}^P$ (the corresponding RIF composition effect). The latter captures the extent to which we failed in reproducing the distribution of characteristics in 1 and is expected to be small if the specification is rich enough:

$$\Delta_S^P = \bar{X}_1\beta_1 - \bar{X}_R\beta_\mathrm{R}$$

$$\Delta_{S,p}^P = \bar{X}_1(\beta_1 - \beta_R);\ \Delta_{S,e}^P = (\bar{X}_1 - \bar{X}_R)\beta_R$$

Similarly, a RIF poverty decomposition of the difference between the reweighted counterfactual and 0 distributions can break the aggregate composition effect into a pure RIF composition effect $\Delta_{X,p}^P$ (the RIF composition term in this decomposition) and a specification error $\Delta_{X,e}^P$ (the corresponding RIF structure term), which takes into account the fact that a change in an average characteristic may also affect conditional poverty:

$$\Delta_X^P = \bar{X}_R\beta_\mathrm{R} - \bar{X}_0\beta_0$$

[11] The detailed decomposition of the composition effect using reweighting can also be estimated by computing the rescaling factors in a sequence of regressions adding a new factor each, or to address omitted variable bias, one single regression but sequentially switching on each coefficient (starting with all set to zero). In both cases there is a problem of path dependence on the order in which factors are accounted for, that can be addressed by applying a Shapley decomposition over all possible sequences (e.g., Gradín 2014). This can be cumbersome with many factors and will not yet identify the detailed decomposition of the structure effect.

$$\Delta_{X,p}^{P} = (\bar{X}_R - \bar{X}_0)\beta_0;\ \Delta_{X,e}^{P} = \bar{X}_R(\beta_R - \beta_0)$$

The corresponding detailed effects can be obtained directly from these terms as in the one-stage RIF decomposition. The analysis can be focused on the detailed effects of interest, that is, the pure RIF composition ($\Delta_{X,p}^{P}$) and pure RIF structure ($\Delta_{S,p}^{P}$) effects. Note that these RIF effects add up to the corresponding reweighting totals only if each error term is zero, something that one would expect in the case of reweighting, but not necessarily in the case of the specification error, which can be case-specific.[12]

6 Concluding remarks

I have shown that RIF regression decompositions can be a very useful tool for understanding what factors are associated with distributional differences over time or between population groups, and poverty is no exception to this. The most common poverty regression-based decomposition found in the empirical literature has estimated these decompositions using regressions of FGT poverty functions, mainly for the headcount ratio, on household characteristics, typically using non-linear models. I have argued that these can be seen as a specific case of the one-stage RIF approach because the poverty functions and the RIF are equivalent concepts in the case of additively decomposable indices. The intuition is that the impact of a marginal increase in the population on a particular income (RIF) will be given by the corresponding individual poverty function, without affecting the poverty functions of other incomes. However, this is not true in the case of rank-based indices such as Sen and extensions, because in those cases, the marginal increase in the proportion of people with an income will affect the rank of the other people and therefore, their poverty functions. In the case of these indices, a regression of the poverty function will be different from a regression of the RIF. However, note that the latter is probably a better representation of the individual impact on overall poverty because it considers both direct and indirect effects, while the former only represents the direct. Furthermore, the combined reweighting/RIF approach (two stages) allows not to impose any functional form on the relationship between characteristics and poverty in obtaining the aggregate decomposition, regardless of the index used. And this is done in a way that is consistent with the method used in the analysis of other distributive measures, since the RIF applies to any index for which the RIF exists. Therefore, the RIF emerges as a more general and consistent approach that is applied to any index and relaxes the need to impose a functional form in the relationship between poverty and characteristics at the aggregate level. Assumptions about the functional form are still necessary to obtain the detailed decomposition using RIF, but the method provides a way to test for a specification error.

References

Blinder, A.S. (1973). 'Wage Discrimination: Reduced Form and Structural Estimates'. *Journal of Human Resources* 8(4): 436-455. https://doi.org/10.2307/144855

Breen, R., K.B. Karlson, and A. Holm (2018). 'Interpreting and Understanding Logits, Probits, and Other Nonlinear Probability Models'. *The Annual Review of Sociology*. https://doi.org/10.1146/annurev-soc-073117-041429

[12] Sampling weights should be used during the entire process. Standard errors can be obtained by bootstrapping the entire process, taking the survey design into account.

Coudouel, A., J.S. Hentschel, and Q.T. Wodon (2002). 'Using Linear Regressions for Analysing the Determinants of Poverty'. In J. Klugman (ed.), *A Sourcebook for Poverty Reduction Strategies*, Annex A.8, Washington DC: World Bank.

DiNardo, J., N.M. Fortin, and T. Lemieux (1996). 'Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach'. *Econometrica* 64: 1001-1044. https://doi.org/10.2307/2171954

Essama-Nssah, B. and P.J. Lambert (2012). 'Influence Functions for Policy Impact Analysis'. In J.A. Bishop and R. Salas (eds), *Inequality, Mobility and Segregation: Essays in Honor of Jacques Silber*, pp. 135–59. Bingley: Emerald Group Publishing Ltd. https://doi.org/10.1108/S1049-2585(2012)0000020009

Even, W.E. and D.A. Macpherson (1990). 'Plant Size and the Decline of Unionism'. *Economics Letters*, 32(4): 393-398. https://doi.org/10.1016/0165-1765(90)90035-Y

Even, W.E. and D.A. Macpherson (1993). 'The Decline of Private-Sector Unionism and the Gender Wage Gap'. *The Journal of Human Resources,* 28(2): 279-296. https://doi.org/10.2307/146204

Fairlie, R.W. (1999). 'The Absence of the African-American Owned Business: An Analysis of the Dynamics of Self-Employment'. *Journal of Labor Economics,* 17(1): 80-108. https://doi.org/10.1086/209914

Fairlie, R.W. (2005). 'An Extension of the Blinder-Oaxaca Decomposition Technique to Logit and Probit Models'. *Journal of Economic Social Measurement,* 30(4): 305–316. https://10.3233/JEM-2005-0259

Firpo, S., N.M. Fortin, and T. Lemieux (2007). 'Decomposing Wage Distributions using Recentered Influence Function Regressions'. Unpublished manuscript, Vancouver: University of British Columbia.

Firpo, S., N.M. Fortin, and T. Lemieux. (2009). 'Unconditional Quantile Regressions'. *Econometrica*, 77(3): 953–73. https://doi.org/10.3982/ECTA6822

Firpo, S., N.M. Fortin, and T. Lemieux (2018). 'Decomposing Wage Distributions Using Recentered Influence Function Regressions'. *Econometrics*, 6(2): 1–40. https://doi.org/10.3390/econometrics6020028

Fortin, N.M., T. Lemieux, and S. Firpo (2011), 'Decomposition Methods in Economics'. In O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*. Amsterdam: North Holland. https://doi.org/10.1016/S0169-7218(11)00407-2

Foster, J.E. (2006). 'Poverty Indices'. In A. de Janvry, R. Kanbur (eds) *Poverty, Inequality and Development. Economic Studies in Inequality, Social Exclusion and Wellbeing*, vol 1. Boston: Springer. https://doi.org/10.1007/0-387-29748-0_4

Foster, J., J. Greer, and E. Thorbecke (1984). 'A Class of Decomposable Poverty Measures'. *Econometrica* 52(3): 761-766. https://doi.org/10.2307/1913475

Gomulka, J., and N. Stern (1990). 'The Employment of Married Women in the United Kingdom 1970-83'. *Economica,* 57: 171–199. https://doi.org/10.2307/2554159

Gradín, C. (2009). 'Why is poverty so high among Afro-Brazilians? A decomposition analysis of the Racial Poverty Gap'. *Journal of Development Studies,* 45(9): 1-38. https://doi.org/10.1080/00220380902890235

Gradín, C. (2014). 'Race and income Distribution: Evidence from the US, Brazil and South Africa'. *Review of Development Economics*, 18(1): 73-92. https://doi.org/10.1111/rode.12070

Gradín, C. (2020). 'Quantifying the Contribution of a Subpopulation to Inequality an Application to Mozambique'. *Journal of Economic Inequality*, 18(3): 391-419. https://doi.org/10.1007/s10888-020-09451-w

Hampel, F.R. (1974). 'The Influence Curve and Its Role in Robust Estimation'. *Journal of the American Statistical Association*, 60: 383-393. https://doi.org/10.1080/01621459.1974.10482962

Kitagawa, E.M. (1955). 'Components of a Difference Between Two Rates'. *Journal of the American Statistical Association,* 50(272): 1168-1194. https://doi.org/10.2307/2281213

Jann, B. (2008). 'A Stata Implementation of the Blinder-Oaxaca Decomposition'. *The Stata Journal*, 8(4): 453-479. https://doi.org/10.1177/1536867X0800800401

Oaxaca, R.L. (1973). 'Male-Female Wage Differentials in Urban Labor Markets'. *International Economic Review,* 14(3): 693-709. https://doi.org/10.2307/2525981

Oaxaca, R.L. and M.R. Ransom (1999). 'Identification in Detailed Wage Decompositions'. *Review of Economics and Statistics,* 81: 154-157. https://doi.org/10.1162/003465399767923908

Schwiebert, J. (2015). 'A Detailed Decomposition for Nonlinear Econometric Models'. *Journal of Economic Inequality* 13: 53-67. https://doi:10.1007/s10888-014-9291-x

Sen, A.K. (1976). 'Poverty: An Ordinal Approach to Measurement'. *Econometrica*, 44: 219-231. https://doi.org/10.2307/1912718

Shorrocks, A.F. (2013). 'Decomposition procedures for Distributional Analysis: A Unified Framework Based on the Shapley Value'. *Journal of Economic Inequality*, 11: 99-126. https://doi.org/10.1007/s10888-011-9214-z

Watts, H.W. (1969). 'An Economic Definition of Poverty'. In D.P. Moynihan (ed.), *On Understanding Poverty*. New York: Basic Books.

Yun, M.-S. (2004). 'Decomposing Differences in the First Moment'. *Economics Letters,* 82(2): 275-280. https://doi.org/10.1016/j.econlet.2003.09.008

Zheng, B. (1997). 'Aggregate Poverty Measures'. *Journal of Economic Surveys,* 11(2): 123-162. https://doi.org/10.1111/1467-6419.00028