Journal of Health Economics 48 (2016) 89–106

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

A general method for decomposing the causes of socioeconomic inequality in health

Gawain Heckley[a,b,*], Ulf-G. Gerdtham[a,b,c], Gustav Kjellsson[d,e]

[a] Health Economics & Management, Institute of Economic Research, Lund University, Box 117, 22100, Sweden
[b] Health Economics Unit, Department of Clinical Sciences, Medicon Village, Lund University, Lund, Sweden
[c] Department of Economics, Lund University, Lund, Sweden
[d] Department of Economics, University of Gothenburg, Gothenburg, Sweden
[e] Centre of Health Economics at Gothenburg University (CHEGU), Gothenburg, Sweden

ARTICLE INFO

Article history:
Received 7 April 2015
Received in revised form 13 March 2016
Accepted 28 March 2016
Available online 7 April 2016

JEL classification:
I10
I14
I30

Keywords:
Inequality measurement
Concentration index
Decomposition methods
Recentered influence function

ABSTRACT

We introduce a general decomposition method applicable to all forms of bivariate rank dependent indices of socioeconomic inequality in health, including the concentration index. The technique is based on recentered influence function regression and requires only the application of OLS to a transformed variable with similar interpretation. Our method requires few identifying assumptions to yield valid estimates in most common empirical applications, unlike current methods favoured in the literature. Using the Swedish Twin Registry and a within twin pair fixed effects identification strategy, our new method finds no evidence of a causal effect of education on income-related health inequality.

1. Introduction

Socioeconomic differences in health are well documented across the western world (Deaton, 2003; Mackenbach et al., 2008, 2015). This awareness has led to a rapidly growing interest in the measurement and analysis of socioeconomic inequality in health. In terms of measurement, the dominant family of measures of socioeconomic inequalities in health are the various versions of the concentration index (CI) – a family of bivariate rank dependent indices. A bivariate rank dependent index summarises the relationship between cumulative health and socioeconomic rank, where a positive or negative socioeconomic gradient in health is represented by a positive or negative index value (Wagstaff et al., 1991; Fleurbaey and Schokkaert, 2009). These measures are bivariate because they relate an individual's level of health to her relative socioeconomic status. They are rank dependent because relative socioeconomic status is given by the socioeconomic rank of the individual.

Policymakers' and researchers' interest in socioeconomic inequality in health also extends beyond measurement through to explaining and understanding its underlying causes. One way to examine this issue is to decompose an inequality measure into a function of its (potential) causes. The dominant decomposition procedure to decompose a bivariate rank dependent index is the technique developed by Wagstaff et al. (2003) (WDW, onwards) which has been used extensively to explore the determinants of the well documented socioeconomic gradient (see, e.g., Leu and Schellhorn, 2004; Gomez and Lopez-Nicholas, 2005; Lauridsen et al., 2007; Hosseinpoor et al., 2006; McGrail et al., 2009; Morasae et al., 2012).[1] As well as being extensively applied, the WDW decomposition method has also been developed to expand its potential for application to a greater set of empirical situations such as health variables that are non-linear in nature (see, e.g.,

* Corresponding author at: Health Economics Unit, Lund University, Medicon Village, SE-223 81 Lund, Sweden. Tel.: +46 766-486666.
E-mail address: Gawain.heckley@med.lu.se (G. Heckley).

[1] Gravelle (2003) is acknowledged for developing the same method although the explicit aim of his paper was not to decompose, but to standardise, the concentration index. The resulting methodology is nevertheless the same as that of WDW decomposition.

http://dx.doi.org/10.1016/j.jhealeco.2016.03.006

Van Doorslaer et al., 2004a,b; Van Ourti et al., 2009; Van de Poel et al., 2009) and the inclusion of heterogeneous responses (Jones and Nicolás, 2006).

The health inequality toolbox is to a large extent adopted from the income inequality literature. The concentration index, for instance, is an adaptation of the Gini index, a popular index in the income inequality literature that measures the degree of income concentration. One important dimension in which measures of socioeconomic related health inequality differ from measures of income inequality, such as the variance or the Gini index, is that the latter consider a single distribution whereas the former consider the joint distribution of health and socioeconomic rank. Specifically to the issue at hand, bivariate rank dependent indices should be thought of as *two-dimensional* indices that consider the *covariance* between health and rank. Unfortunately, the leading decomposition method for bivariate rank dependent indices, the WDW decomposition method, is one-dimensional because it focuses on health but ignores rank (Erreygers and Kessels, 2013). That is, the WDW decomposition method explains the degree of variation in health rather than the covariance between health and rank. In response to this, Erreygers and Kessels (2013) and Kessels and Erreygers (2015) derive a set of two-dimensional decomposition methods, where rank and health are both estimated.[2] However, both the WDW decomposition and the alternatives proposed by Erreygers and Kessels (2013) and Kessels and Erreygers (2015) also suffer from another limitation: they can only *correctly* decompose one form of rank dependent index, yet there is an abundance of rank dependent measures used in the literature. So whilst one may wish to measure inequality in different ways, these approaches will yield the same results no matter which measure one chooses. We will illustrate later how the WDW method is only able to decompose absolute inequality measures (i.e., measures invariant to the addition or subtraction of an equal amount of health for all individuals in the population), which for example does not include the standard concentration index.

The various issues with the WDW methodology are closely linked to the general critique of decomposition methods raised by Fortin et al. (2011): many decomposition methods have focused on the derivation of procedures without first specifying the object of interest nor how to identify this object (i.e., stating what we want to estimate and the assumptions required to interpret the estimates). Indeed, both the derivation of the WDW method and the subsequent variations thereof have focussed on the procedure rather than identification of the object of interest. The discussion of identification has come second at best, potentially because these decompositions have been seen as an accounting exercise. As a consequence it is unclear how to actually interpret the parameters, and the so called contributions, within these decompositions. The relevance of this critique for the existing variations of the WDW decomposition is implicitly illustrated by Erreygers and Kessels (2013). Following a similar line of logic to that of the standard WDW decomposition, Erreygers and Kessels (2013) derived a set of two-dimensional decompositions by making small changes to the starting point of the procedure. These different methods yield a wide range of results, yet it is unclear which is preferred, and how to interpret the estimated coefficients. Erreygers and Kessels (2013) were unable to choose a preferred method because, as they noted themselves, they did not consider the identification issue.[3] In sum, the literature has shown that the WDW method is not only a one-dimensional decomposition of just one specific form of a bivariate rank dependent index but also just one possible method of many similar alternatives, the choice of which greatly affects the results.

We contribute to this literature by deriving and empirically illustrating an alternative regression based decomposition method for rank dependent indices that overcomes the criticisms of the decomposition methods currently available. This method aims to explain the causes of socioeconomic inequality, not by focussing on the variables that form the covariance, but by directly decomposing the weighted covariance of health and socioeconomic rank, i.e., the rank dependent index. This new approach builds on the concept of regression of a recentered influence function (RIF). A RIF is a concept that originates from the robustness literature of statistics that yields an approximation of the derivative (gradient) of a statistic. Intuitively, the RIF is a vector where each element corresponds to a particular individual's influence on the statistic. That is the RIF value for a specific individual tells us how the statistic would change if that individual were to be removed from the sample (weighted by the inverse of the sample size). The RIF is useful for decomposition because it allows any statistic to be expressed as a mean of the RIF vector and this allows all the regression tools for standard mean analysis to be used to link individual characteristics to a statistic. Importantly RIF regression already has a defined object of interest and it is clear how to identify it (Firpo et al., 2009): regressing the vector of RIF values on a set of covariates yields the unconditional partial effect of covariates on the statistic.

In this paper we apply the concept of RIF regression to a bivariate rank dependent index. Firpo et al. (2007, 2009) applied RIF regression to an income inequality question, estimating and decomposing RIFs for univariate measures such as the variance, the unconditional quantile, and the Gini index. The major contribution of this paper, and a key step forward for health inequality analysis, is that we derive the RIF for a general bivariate rank dependent index, and also specifically for familiar versions such as the concentration index and the adjustments suggested by Wagstaff (2005) and Erreygers (2009). Decomposition of the index is then performed by a two-step procedure of first computing the RIF of the rank dependent index, and then regressing the RIF on a set of covariates yielding the marginal effects of the covariates on the index.

The application of the RIF regression method to the decomposition of rank dependent indices has a few important benefits. First, the object of interest of the method is clear and therefore discussion of identification is much more straightforward. Second, the method directly decomposes the weighted covariance of health and socioeconomic rank. As a consequence it overcomes the critique of Erreygers and Kessels (2013) and it can be used to decompose all forms of bivariate rank dependent indices. Being able to decompose all forms of bivariate rank dependent indices is a key feature of this new method because each form of rank dependent index has a different set of underlying value judgements with respect to inequality (Allanson and Petrie, 2014; Kjellsson et al., 2015) and there remains no actual consensus as to which index is preferred. The ability to decompose several indices is therefore key for health inequality analysis.

A further benefit of RIF regression decomposition is that the results are familiar in their interpretation. Assuming a linear relationship means the RIF is the dependent variable in an OLS regression whose coefficients equal the marginal effect of covariates X on the rank dependent index. This interpretation is analogous to that of an OLS regression of a random variable. Indeed, a RIF decomposition of the

[2] In the more recent paper by Kessels and Erreygers (2015), they propose a structural equation modelling approach where rank and health are both estimated. This two-dimensional decomposition is one potential way to acknowledge the bivariate nature of these inequality indices, but the requirements of such a structural modelling approach is data demanding (requires two instrumental variables for health and for rank respectively), limiting the potential scope for such a solution. This solution also doesn't address the other issues raised in this paper.

[3] In the conclusion of their paper, Erreygers and Kessels (2013) call for an "axiomatic approach" to derive the most preferred method. We interpret this as a call for specifying the object of interest in beforehand, and then set up the assumptions needed to identify this object.

mean (assuming a linear function of the dependent variable) is simply OLS of a random variable (Firpo et al., 2009). As most researchers are familiar with OLS, assuming linearity makes the RIF regression straightforward to estimate and the conditions needed to obtain a causal parameter are well known. This familiarity makes RIF decomposition a useful tool, not only for descriptive analysis, but also in a policy evaluation framework.

In order to help the reader understand why RIF regression based decomposition is a useful addition to the analyst's toolkit for the analysis of bivariate rank dependent indices, a brief description of rank dependent indices and the standard WDW decomposition method is provided (in Section 2) before a discussion of the identifying assumptions of WDW decomposition (Section 2 again). Although the health inequality literature has previously highlighted that these identifying assumptions may be restrictive (Van Doorslaer et al., 2004a,b; Erreygers and Kessels, 2013; Gerdtham et al., 2016), they have never been summarised clearly in one place, and may therefore be unknown to practitioners. The literature suggests that the usefulness of the WDW decomposition should be questioned, as the violation of these conditions is potentially severe. The paper then presents a new method for decomposing a bivariate rank dependent index based upon RIF regression that requires fewer identifying assumptions. To help develop the intuition of this new method the concept of the RIF is briefly introduced before deriving the RIF for a general bivariate rank dependent index (Section 3). RIF decomposition is then discussed in detail (Section 4).

To illustrate the differences in interpretation between the RIF and the WDW decomposition we present an empirical example using the Swedish Twin Register (Section 5). The empirical example also highlights the importance of being able to decompose different forms of rank dependent index showing that the choice of index has bearing on the association between education and health inequality. We find no association of education with socioeconomic health inequality using RIF regression. To highlight how one can use the RIF decomposition for establishing causal relationships, we use a twin differencing strategy to attempt to isolate the effect of education on socioeconomic related heath inequality. The results suggest there is no causal effect of education on any common choice of bivariate rank based measure of health inequality.

Having illustrated RIF regression of a bivariate index we then discuss the relative merits of this new approach compared to WDW decomposition (Section 6) concluding that RIF-I-OLS will uncover the (causal) parameters of interest under common empirical conditions. Evidence from the literature and also presented in this paper suggests that when concern lies with covariates that are known to impact on the ranking variable and or the weighting variable, WDW decomposition is likely to yield biased results. Conversely, RIF regression does not require these identifying assumptions and this makes the RIF regression of a bivariate rank dependent index easier to interpret and a preferable descriptive decomposition tool. In addition, RIF regression is also well suited to policy evaluation. RIF regression allows the effect of a policy to be evaluated across a wide range of statistics, highlighting its potential in the field of program evaluation.

2. Preliminaries

2.1. A rank dependent index

The general term for a statistic, such as the mean, variance or the Gini for example, is a functional, $v(F)$, where F is a probability measure for which $v(F)$ is defined.[4] Let us define $H \in [0,+\infty)$[5] as a random variable of health with mean denoted as μ_H and with probability measure denoted as F_H. We rank each individual by a random variable for socioeconomic status, Y. The CDF of Y, F_Y, yields the fractional rank for each individual, which by definition has mean ½ (F_Y is uniformly distributed over the unit interval). The joint distribution of H and F_Y is given by F_{H,F_Y}. The functional for the general form of a rank dependent index (I) is then given by:

$$I = v^I(F_{H,F_Y}) = v^{\omega_I}(F_H) v^{AC}(F_{H,F_Y}), \quad (1)$$

where $v^{\omega_I}(F_H)$ is a weighting function specific to a particular form of rank dependent index, and the absolute concentration index (AC) is given by twice the covariance between H and F_Y:

$$AC = v^{AC}(F_{H,F_Y}) = 2cov(H, F_Y), \quad (2)$$

We refer to this as the absolute concentration index as it is invariant to the addition or subtraction of an equal amount of health for all individuals in the population.[6] The relative counterpart is the standard concentration index (CI), which is invariant to equi-proportional changes in health. The weighting functions for these common forms of rank dependent index are:

Absolute concentration index:

$$v^{\omega_{AC}}(F_H) = 1 \quad (3)$$

Concentration index:

$$v^{\omega_{CI}}(F_H) = \frac{1}{\mu_H} \quad (4)$$

Different choices of weighting function imply different value judgements, in this case a preference for absolute or relative inequality. The choice of index, and therefore the choice of weighting function, is more complex when the health variable of interest has both an upper and lower bound denoted as b_H and as a_H respectively, i.e., $H \in [a_H, b_H]$ (Wagstaff, 2005; Erreygers, 2009; Erreygers and Van Ourti, 2011; Kjellsson and Gerdtham, 2013a,b; Kjellsson et al., 2015). For such a variable, health can be represented as both attainments $(H - a_H)$ and shortfalls $(b_H - H)$, and the choice of which affects the value of the concentration index. One set of indices adjusted for bounded variables

[4] The symbol v is used to signify a functional and comes from the class of statistics called v-statistics.

[5] We define H in the general case as an unbounded measure without any loss of generalisability for bounded health variables.

[6] In the literature the absolute concentration index is sometimes called the generalised concentration index, although it is not a generalisation of the concentration index. We label it the absolute concentration index because it is an absolute measure of socioeconomic-related health inequality (it is not affected by the addition or subtraction of a certain amount of health).

assures that the level of inequality is the same irrespective of this representation. The weighting functions for two rank dependent indices that make this adjustment are:

Erreygers index:

$$\nu^{\omega_{EI}}(F_H) = \frac{4}{b_H - a_H} \tag{5}$$

Wagstaff index:

$$\nu^{\omega_{WI}}(F_H) = \frac{b_H - a_H}{(b_H - \mu_H)(\mu_H - a_H)} \tag{6}$$

The Erreygers Index (EI) is an absolute index adjusted for a bounded variable, whereas the underlying value judgement of the Wagstaff Index (WI) is more complex (Wagstaff, 2005; Kjellsson and Gerdtham, 2013a,b; Allanson and Petrie, 2014). It is also possible to define a concentration index that is invariant to either proportional changes in attainment or shortfalls of bounded health variables. Following Kjellsson et al. (2015), we denote these as:

Attainment-relative concentration index (ARCI)[7]

$$\nu^{\omega_{ARCI}}(F_H) = \frac{1}{(\mu_H - a_H)} \tag{7}$$

Shortfall-relative concentration index[8] (SRCI)

$$\nu^{\omega_{SRCI}}(F_H) = \frac{1}{(b_H - \mu_H)} \tag{8}$$

There exists no actual consensus as to which index is preferred, but the literature stresses that any choice of index represents a value judgement (Allanson and Petrie, 2014; Kjellsson et al., 2015). Given this lack of consensus it is arguably important that any decomposition analysis is able to encompass as broad a view as possible.

2.2. The standard decomposition

The leading decomposition method applied to I is the WDW decomposition method based on a linear regression of health. Assuming health, represented by h, an $n \times 1$ vector of drawings from H, is observed alongside covariates, X, and that health can be expressed as a linear in variables model in X, together yields the following regression equation:

$$h = \alpha + X'\beta + e, \tag{9}$$

where X is a $k \times n$ matrix, α is an intercept, β is a $k \times 1$ vector of regression coefficients, and e is a $n \times 1$ vector of error terms. Following Wagstaff et al. (2003), I can then be decomposed by substituting Eq. (9) into (1), yielding the following formula:

$$I = \nu^I(F_{H,F_Y}) = \nu^{\omega_I}(F_H)\sum_{k=1}^{K}\beta_k 2cov(X_k, F_Y) + \nu^{\omega_I}(F_H) 2cov(e, F_Y), \tag{10}$$

where β_k is the regression coefficient corresponding to the kth regressor from the linear regression Eq. (9), $2cov(X_k, F_Y)$ is the absolute concentration index of the kth covariate X_k and $2cov(e, F_y)$ is the absolute concentration index of e. The first part of the WDW decomposition formula, given by Eq. (10), expresses the change in $\nu^I(F_{H,F_Y})$ predicted by a change in either $cov(X_k, F_Y)$ or β_k, what we will call *marginal contributions*. The first part of equation (10) has also been used to express I as the proportion explained by X, "the explained part" (what we will refer to as *percentagewise contributions*), plus the second part of Eq. (10), as "the unexplained part". What can be immediately observed from Eq. (10) is that the WDW decomposition implicitly holds the weighting function, $\nu^{\omega_I}(F_H)$, constant and assumes that if one made any changes that impacted $\sum_{k=1}^{K}\beta_k 2cov(X_k, F_Y)$ this would have no impact on the weighting function. A consequence of this is that percentagewise contributions are the same no-matter which index one uses.

2.3. The identifying assumptions of WDW decomposition

To the best of our knowledge the identifying assumptions that underpin WDW decomposition, whilst not new to the literature, have never been stated explicitly in one place, neither in its application or otherwise. To be explicit we set out these assumptions below and then discuss each assumption in turn.

The identifying assumptions required by the WDW decomposition are:

I. The determinants of health do not determine rank (rank ignorability).
II. The determinants of health do not determine the weighting function (weighting function ignorability).
III. Health can be modelled as a function linear in variables X and an error term.
IV. Exogeneity: The errors from the health regression have zero conditional mean.

[7] Erreygers and Van Ourti (2011) first suggested this as the generalised version of the corrected concentration index.

[8] An index using this weighting function is equivalent to applying the (attainment-relative) concentration index representing the health variable in terms of shortfalls, or ill health.

If all the identifying assumptions above hold, WDW decomposition identifies both percentagewise and marginal contributions yielding results of potentially great empirical interest.[9] In most empirical applications identifying assumption IV – which OLS requires for causal interpretation – is not seen as a necessary condition and WDW decomposition is generally viewed as a "simple descriptive accounting exercise" based on some correlations from an OLS regression (Gerdtham et al., 2016). WDW decomposition is therefore generally thought of as yielding descriptive percentagewise contributions. However, even as a descriptive accounting exercise, this still requires the results to be interpreted in light of identifying assumptions I, II & III, which in empirical practice often are unreasonable to impose. This muddies the interpretation of the results.

The restrictiveness of *rank ignorability* (Identifying assumption I) has previously been pointed out by Erreygers and Kessels (2013) as well as in Kessels and Erreygers (2015). They criticise the WDW decomposition approach for being a *one-dimensional* decomposition (of a bivariate index) because it only decomposes one part of the covariance (health). Ignoring the association between the covariates and rank means that for any (causal) explanation of changes in covariates the income rank is assumed to remain the same even after the change. Indeed Erreygers and Kessels (2013) and Kessels and Erreygers (2015) both find important differences in the results when this assumption is relaxed. It is important to note that their results are based on an approach that still maintains the other identifying assumptions.

Assumption II, *weighting function ignorability*, is similarly restrictive because the weighting function, $\nu^{\omega_I}(F_H)$, is generally a function of health, and will by design be correlated with the covariates, as the covariates are predictors of health. As seen in Eqs. (3)–(8), the weighting functions of CI, WI, ARCI, and SRCI are all functions of mean health. Only absolute versions of the rank dependent index (such as AC and EI) have a constant weighting function. Because w*eighting function ignorability* requires the analyst to assume that the weighting function is unaffected by a change in these covariates, a WDW decomposition of any rank dependent index implicitly decomposes an absolute version of the index. In practice this restriction means that the WDW decomposition is only applicable to absolute inequality indices (even though it was developed for the relative concentration index).[10]

In regard to assumption III, *health is a function linear in variables*; there are few health outcomes that can truly be modelled in a linear way. It is common to find non-linear health functions: outcomes may be categorical (Underweight, normal, overweight, obese), censored at zero (doctor visits) or two-part decisions (quantity smoked) all of which are non-linear. The linearity assumption of WDW, however, requires more than the standard linearity assumption: To provide the popular interpretation of percentagewise contributions of each variable of interest, WDW requires the model to be not only linear in parameters, but linear in variables. Potential solutions have been proposed (see, e.g., Van Doorslaer et al., 2004a,b; Van Ourti, 2004; Van de Poel et al., 2009; Van Ourti et al., 2009), but they require the non-linear estimates to be translated back to the linear setting in order to yield percentagewise contributions or are applicable only to changes. With that said, the linearity assumption has not been found to be that restrictive in practice (Van Doorslaer et al., 2004b; Van de Poel et al., 2009). In addition, linearity is an assumption that empirical economists are often willing to make (especially in the policy evaluation literature). The flexibility and simplicity of methods such as OLS generally provide a powerful framework for empirical analysis. The available evidence does indicate that assumption III ranks as a less restrictive assumption compared to both I and II.

The underlying issue with the current available methods to decompose a bivariate rank dependent index is the critique raised by Fortin et al. (2011): many decomposition methods have not been explicit about what the parameter of interest is and the required identifying assumptions. The methods that are currently available have been developed with a focus on procedures with little thought given to identification. As set out in the introduction, Erreygers and Kessels (2013) implicitly illustrate the consequence of this ambiguity and derive quite a few alternative decomposition methods. A consequence of not defining the parameter of interest is that it is not immediately obvious how to interpret these various different methods of decomposition yet alone be able to choose a preferred method. However, the results obtained by Erreygers and Kessels (2013) vary quite dramatically depending on the method chosen and therefore the choice of method matters. Combined with the literature highlighting the implicit identifying assumptions of WDW, the findings of Erreygers and Kessels (2013) reveal that the results of the WDW decomposition are not as easily interpreted as once thought.

In the next section we derive a completely different approach to regression-based decomposition of a bivariate rank dependent index that allows two of the identifying assumptions of WDW decomposition to be relaxed simultaneously: rank and weighting function ignorability. Importantly, we explicitly state our parameter of interest and the assumptions required to identify this parameter. This method has the potential to identify the parameters of interest under much more common empirical conditions, yielding results that have a clear interpretation.

3. The RIF for a general bivariate rank dependent index

The RIF is derived from the influence function (IF), which originates from the robustness literature of statistics. Hampel (1974) introduced the concept of the IF with the original purpose to explore how various statistics are affected (or influenced) by particular observations, hence the name, influence function. The RIF has the same properties as the IF with the singular exception that the RIF has a different expected value to that of the IF. Firpo et al. (2009) developed the concept of the RIF, RIF regression and hence RIF decomposition. In this section we first introduce the concept of the IF and the RIF in a univariate setting, before deriving the RIF for a general *bivariate* rank dependent index.

3.1. The influence function and the recentered influence function

The influence function is a specific form of a directional derivative (or Gâteaux derivative). A directional derivative is used to find the influence of a perturbation or contamination in a distribution, for example from F_H towards a new distribution, on a statistic. The IF is

[9] It is worth noting that percentagewise contributions is a global parameter and requires the assumption that under large changes in the covariate there will be no (unaccounted for) general equilibrium effects.

[10] A Potential solution to weighting function ignorability has been developed by Van Ourti et al. (2009) but this still assumes rank ignorability and is applicable only for decomposition of changes.

the particular form of a directional derivative where the new distribution, denoted as δ_h, equals a cumulative distribution function for a probability measure that puts mass 1 at a particular value *h:*

$$\delta_h(l) = \begin{cases} 0 & \text{if } l < h \\ 1 & \text{if } l \geq h \end{cases}, \tag{11}$$

where *l* is a draw from *H*.[11] To define the IF of the functional $v(F_H)$ evaluated at point *h*, denoted as IF(h; v), we first define G_h as a mixing probability distribution of F_H and δ_h:

$$G_h = (1 - \varepsilon)F_H + \varepsilon\delta_h, \tag{12}$$

where $\varepsilon \in (0, 1)$ is a probability, or a weight, representing the relative change in the population through the addition of δ_h. That is, G_h is a distribution that is ε away from F_H in the direction of δ_h. IF(h; v) is then defined as:

$$IF(h; v) = \left.\frac{\partial v(G_h)}{\partial \varepsilon}\right|_{\varepsilon=0} = lim_{\varepsilon\to 0}\frac{v(G_h) - v(F_H)}{\varepsilon}, \tag{13}$$

if the limit is defined for every point $h \in \mathbb{R}$, where $\mathbb{R}$ is the real line.[12] Intuitively speaking, the IF captures the (limiting) influence of an *individual* observation on the functional $v(F_H)$ (Wilcox, 2005) and this can be used to understand how the addition/subtraction of an observation would affect a statistic without having to re-calculate the statistic. In practice, calculating an IF yields an influence function value for each individual in the sample.

Having defined the IF it is now possible to define the RIF. One can think of the RIF in two ways. First, as a linear approximation of the functional, the RIF consists of the first two leading terms of a Von Mises linear approximation. The RIF is also a minor transformation of the IF, and is obtained from the IF by adding back the original functional, $v(F_H)$:

$$RIF(h; v) = v(F_H) + IF(h; v). \tag{14}$$

While the expectation of the IF is zero (Monti, 1991), the expectation of the RIF is equal to the original distributional statistic $v(F_H)$ (Firpo et al., 2009). This is a useful property because, as we discuss later, it allows standard regression tools for the mean to be applied to (and therefore decompose) any statistic.

To illustrate the two concepts, the IF and the RIF, assume the statistic of interest is the mean. The IF of μ_H equals IF(h ; μ_H) = $lim_{\varepsilon\to 0}((1-\varepsilon)\mu_H + \varepsilon h - \mu_H)/\varepsilon = h - \mu_H$. This states that adding or removing an observation will have an effect on μ_H equal to the distance between the observation, *h*, and the mean (standardised by the sample size). Adding the statistic, μ_H, to $IF(h\,;\mu_H)$ yields the RIF of the mean, $RIF(h\,;\mu_H) = \mu_H + (h - \mu_H) = h$.

3.2. The RIF for a general (bivariate) rank dependent index

As the rank dependent index, *I*, is a functional of the joint probability distribution F_{H,F_Y}, we need to extend the definitions in Eqs. (11)–(14) from a univariate to a bivariate setting. Let $G_{h,F_Y(y)}$ be a bivariate distribution function obtained by an infinitesimal contamination of F_{H,F_Y} in both *h* and $F_Y(y)$:

$$G_{h,F_Y(y)} = (1 - \varepsilon)F_{H,F_Y} + \varepsilon\delta_{h,F_Y(y)}. \tag{15}$$

Here $\delta_{h,F_Y(y)}$ denotes a joint cumulative distribution function for a joint probability measure that gives mass 1 to $(h, F_Y(y))$ jointly:

$$\delta_{h,F_Y(y)}(l, r) = \begin{cases} 0 & \text{if } l < h \text{ or } r < F_Y(y) \\ 1 & \text{if } l \geq h \text{ and } r \geq F_Y(y) \end{cases}, \tag{16}$$

where *l* and *r* are draws from *H* and F_Y respectively. In analogy with Eq. (13), we then define the bivariate IF of $v^I(F_{H,F_Y})$ evaluated at point $(h, F_Y(y))$ as[13]:

$$IF(h, F_Y(y); v^I) = \left.\frac{\partial v^I(G_{h,F_Y(y)})}{\partial \varepsilon}\right|_{\varepsilon=0} = lim_{\varepsilon\to 0}\frac{v^I(G_{h,F_Y(y)}) - v^I(F_{H,F_Y})}{\varepsilon}, \tag{17}$$

given that this limit is defined for every point $(h, F_Y(y)) \in \mathbb{R}^2$, where $\mathbb{R}^2$ denotes the real plane. The RIF of *I* is then defined as:

$$RIF(h, F_Y(y); v^I) = v^I(F_{H,F_Y}) + IF(h, F_Y(y); v^I). \tag{18}$$

In Proposition 1 we state the expression of the RIF for a general bivariate rank dependent index for socioeconomic related health inequality, leaving the proof to Appendix A, before we present the RIF for the common forms of *I* that appear in the health inequality literature.

---

[11] Note that the definition of *h* is no longer an $n \times 1$ vector as defined previously for the WDW decomposition.

[12] Another way of checking whether the IF exists is to check if the functional is continuous (has no jumps or spikes) and the differential is bounded.

[13] Note that *I* is a covariance not of two random variables but a covariance of a random variable, *H*, and the ranking variable, F_Y, which is a function of a random variable. Deriving the IF is therefore more complicated than deriving the IF of a standard covariance because the ranking function is also affected by the infinitesimal contamination.

Proposition 1. *Let $\nu^I(F_{H,F_Y}) = \nu^{\omega_I}(F_H)\nu^{AC}(F_{H,F_Y})$ be a general rank dependent index, the AC be defined as $\nu^{AC}(F_{H,F_Y}) = 2cov(H, F_Y)$ and F_{H,F_Y} be the joint CDF of H and F_Y with corresponding pdf denoted as f_{H,F_Y}. Then the RIF for $\nu^I(F_{H,F_Y})$ is given by:*

$$RIF(h, F_Y(y); \nu^I) = \nu^I(F_{H,F_Y}) + IF(h; \nu^{\omega_I}) * \nu^{AC}(F_{H,F_Y}) + \nu^{\omega_I}(F_H) * IF(h, F_Y(y); \nu^{AC}),$$

where $IF(h; \nu^{\omega_I})$ denotes the IF of the weighting function for I and $IF(h, F_Y(y); \nu^{AC}) = -2\nu^{AC}(F_{H,F_Y}) + \mu_H - h + 2hF_Y(y) - 2\int^y \int^{+\infty} hf_{H,F_Y} dhdF_Y(z)$ denotes the IF for AC.

Proposition 1 shows that for any general rank dependent index, the RIF of I equals the sum of the original statistic, $\nu^I(F_{H,F_Y})$, and its IF, of which the IF is found by application of the product rule of $\nu^{\omega_I}(F_H)\nu^{AC}(F_{H,F_Y})$. The IF for the AC consists of terms familiar from standard inequality analysis; the AC, the mean of health, an individual's health, an individual's rank, and the absolute concentration curve co-ordinate of the individual, $\nu^{ACC}(F_{H,F_Y}(y)) = (2\int^y \int^{+\infty} hf_{H,F_Y} dhdF_Y(z))$. The RIF of any I follows from calculating the IF of the weighting function for the particular I in question and then slotting this into the formula for the RIF given in Proposition 1. Corollary 1 presents the formulas for the RIF of the specific versions of I, again leaving the proof to Appendix A.[14]

Corollary 1. *The RIFs for the AC, EI, CI, ARCI, SRCI and the WI are given by*:

$$RIF(h, F_Y(y); \nu^{AC}) = \nu^{AC}(F_{H,F_Y}) + IF(h, F_Y(y); \nu^{AC})$$

$$RIF(h, F_Y(y); \nu^{EI}) = \nu^{EI}(F_{H,F_Y}) + \frac{4}{b_H - a_H} IF(h, F_Y(y); \nu^{AC})$$

$$RIF(h, F_Y(y); \nu^{CI}) = \nu^{CI}(F_{H,F_Y}) + \frac{(\mu_H - h)}{\mu_H^2} * \nu^{AC}(F_{H,F_Y}) + \frac{1}{\mu_H} IF(h, F_Y(y); \nu^{AC})$$

$$RIF(h, F_Y(y); \nu^{ARCI}) = \nu^{ARCI}(F_{H,F_Y}) + \frac{(\mu_H - h)}{(\mu_H - a_H)^2} * \nu^{AC}(F_{H,F_Y}) + \frac{1}{\mu_H - a_H} IF(h, F_Y(y); \nu^{AC})$$

$$RIF(h, F_Y(y); \nu^{SRCI}) = \nu^{ARCI}(F_{H,F_Y}) + \frac{(-\mu_H + h)}{(b_H - \mu_H)^2} * \nu^{AC}(F_{H,F_Y}) + \frac{1}{b_H - \mu_H} IF(h, F_Y(y); \nu^{AC})$$

$$RIF(h, F_Y(y); \nu^{WI}) = \nu^{WI}(F_{H,F_Y}) + \frac{-(b_H - a_H)[(b_H + a_H - 2\mu_H)(h - \mu_H)]}{((b_H - \mu_H)(\mu_H - a_H))^2} * \nu^{AC}(F_{H,F_Y}) + \frac{b_H - a_H}{(b_H - \mu_H)(\mu_H - a_H)} IF(h, F_Y(y); \nu^{AC})$$

The RIF formulas may appear complex, however they are just a linearisation of the statistic. Practical implementation of RIF estimation is straight forward and to illustrate this we consider empirical estimation of the RIF where the empirical RIF for I is estimated using sample data as:

$$\widehat{RIF}(h, F_Y(y); \nu^I) = \widehat{\nu^I}(F_{H,F_Y}) + \widehat{IF}(h; \omega_I) * \widehat{\nu^{AC}}(F_{H,F_Y}) + \widehat{\nu^{\omega_I}}(F_H)\left[-\widehat{\nu^{AC}}(F_{H,F_Y}) + \widehat{\mu_H} - h_i + 2h_i\widehat{F_Y}(y_i) - \widehat{\nu^{ACC}}(F_{H,F_Y}(y_i))\right] \tag{19}$$

To empirically estimate the RIF, the data of N observations is first ordered using the ranking variable, Y, so that $y_1 \le y_2 \le \ldots \le y_i \le \ldots \le y_N$. Then estimates of the I, $\widehat{\nu^I}(F_{H,F_Y})$, the AC, $\widehat{\nu^{AC}}(F_{H,F_Y})$, the weighting function, $\widehat{\nu^{\omega_I}}(F_H)$, and the mean $\widehat{\mu_H}$ are obtained using the formulas in Section 2. The estimate of the rank, $\widehat{F_Y}(y_i)$, and the absolute concentration curve coordinate, $\widehat{\nu^{ACC}}(F_{H,F_Y}(y_i))$, can be calculated as follows:

$$\widehat{F_Y}(y_i) = \frac{\sum_i^1 1}{N} \tag{20}$$

$$\widehat{\nu^{ACC}}(F_{H,F_Y}(y_i)) = \frac{\sum_i^1 h_i}{N}, \tag{21}$$

where the numerators are a sum that follow the orderings of the i values of Y.[15] Together these yield the empirical RIF. It is important to note that the formulas are the same for all empirical applications, no matter what form the health and socioeconomic ranking variable. Consequently, estimation of the RIF can be automated. To this end the Stata do file used in our empirical example of this paper is provided as an Appendix (see online supplementary material) that allows estimation of the RIFs derived in this paper and also provides a working example of how to decompose the RIF and yield bootstrapped standard errors.[16] For decomposition analysis (RIF regression) the empirical RIF is used as a dependent variable in a regression. We now turn to the concept of RIF regression.

---

[14] The formula for the RIF for the CI is very similar to the RIF for the univariate Gini. Indeed we show in Appendix B that if we derive the RIF for the univariate Gini from the covariance formula, as we have done in the proof of proposition 1, this is the same as presented in Firpo et al. (2007) where the RIF for the Gini has been derived from a formula for the Lorenz curve.

[15] The absolute concentration curve is a mapping of cumulative health and fractional rank (Wagstaff et al., 2003).

[16] A Stata ado file is also available that allows users to perform OLS based RIF regression of a number of forms of the concentration index and also save the RIF values to perform graphical analysis and is found at: https://sites.google.com/site/gawainheckley/home/stata-code?pli=1.

4. RIF regression decomposition

RIF regression is a method that allows us to decompose a RIF of any functional into a function of the sources of its variation, the covariates, X. Our focus is decomposition of I and hence RIF of I regression decomposition. Firpo et al. (2009) identify two parameters of interest that can be estimated using RIF regression: the marginal effect of covariates X on a functional, which is an individual effect, and the unconditional partial effect, which is a population effect measure. The latter captures the impact of a marginal location shift in a continuous covariate or the impact of marginal changes in the conditional distribution of a binary covariate holding everything else constant. Relating this to the topic of the paper, the unconditional partial effect measures how an equal marginal increase in education for everyone would impact on the bivariate rank dependent index (Note that RIF regression estimates *marginal contributions*, not percentagewise contributions as WDW decomposition results are often presented). In this section we show how RIF regression of I obtains these parameters and the assumptions required to identify them.

4.1. RIF regression

The recentering of the IF yielding the RIF implies that $\nu^I(F_{H,F_Y})$ can be expressed as an expected value of the RIF:

$$\nu^I(F_{H,F_Y}) = \int_{-\infty}^{\infty} RIF(h, F_Y(y); \nu^I) \cdot dF_{H,F_Y}(h, F_Y(y)) = E[RIF(H, F_Y; \nu^I)] \tag{22}$$

In order to link $\nu^I(F_{H,F_Y})$ to the covariates X, we follow Firpo et al. (2009) applying the law of iterated expectations to express $\nu^I(F_{H,F_Y})$ as a conditional expectation:

$$\begin{aligned}\nu^I(F_{H,F_Y}) &= \int_{-\infty}^{\infty} RIF(h, F_Y(y); \nu^I) \cdot dF_{H,F_Y}(h, F_Y(y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} RIF(h, F_Y(y); \nu^I) \cdot dF_{(H,F_Y)|X}(h, F_Y(y)|X = x) \cdot dF_X(x) \\ &= \int_{-\infty}^{\infty} E[RIF(H, F_Y; \nu^I)|X = x] \cdot dF_X(x)\end{aligned} \tag{23}$$

where F_X is the CDF of X.[17] Thus, decomposing $\nu^I(F_{H,F_Y})$ boils down to a problem of estimating a conditional expectation, which can be solved by standard regression methods. For a general function of covariates X and an error term $\in$, denoted as $\lambda(X, \in)$, the conditional expectation of $RIF(h, F_Y(y); \nu^I)$ may then be modelled as:

$$E[RIF(H, F_Y; \nu^I)|X = x] = \lambda(X, \in) \tag{24}$$

The first parameter of interest, the marginal effect with respect to X, is given by the partial derivative of the regression estimates of (24):

$$\frac{dE[RIF(H, F_Y; \nu^I)|X = x]}{dx} = \frac{d\lambda(X, \in)}{dx} \tag{25}$$

The second parameter of interest is the unconditional I partial effect. For a continuous covariate, this captures the response of I to a small location shift in the covariate (unconditional on the other covariates). For a binary covariate, this captures the response of I to marginal changes in the conditional distribution of the binary covariate *given* the other covariates. The $k \times 1$ vector of unconditional I partial effects, denoted as $\gamma(\nu^I)$, is a vector of average partial derivatives expressed as:

$$\gamma(\nu^I) = \int_{-\infty}^{\infty} \frac{dE[RIF(H, F_Y; \nu^I)|X = x]}{dx} \cdot dF_X(x) = \int_{-\infty}^{\infty} \frac{d\lambda(X, \in)}{dx} \cdot dF_X(x) \tag{26}$$

The potential choice of regression methods one could use to model the conditional expectation of $RIF(h, F_Y(y); \nu^I)$ and recover these parameters is limitless, but the eventual choice will depend on the form one is willing to assume for the function $\lambda(.)$. Assuming $\lambda(.)$ to be linear and applying OLS to estimate the parameters, yields an estimator we refer to as RIF-I-OLS. We use RIF-I-OLS as our working example for illustration of the method, because it is both simple and attractive from an operational perspective. As is the case for standard OLS, the restriction to a linear in parameters functional form, still allows for a fairly flexible functional form by inclusion of non-linear or higher order transformations of the covariates.

---

[17] Noting that $F_{H,F_Y}(h, F_Y(y)) = \int F_{(H,F_Y)|X}(h, F_Y(y)|X = x) \cdot dF_X(x)$, which is substituted into the second equality.

4.2. RIF-I-OLS

RIF-I-OLS identifies our parameters of interest, the marginal effect and the unconditional I partial effect, under the following assumptions:

Additive linearity. Assuming a functional form linear in parameters with an additive error term for the regression model for the RIF of I, we may rewrite Eq. (24) as:

$$E[RIF(H, F_Y; \nu^I)|X = x] = X'\psi + \mu \tag{27}$$

Zero conditional mean. $E[\mu|X] = 0$. Assuming conditional mean independence of the error term means our coefficient estimates, ψ, have a meaningful interpretation.

Using ν^{AC} as an example, assuming a linear functional form implies the assumption that the sum of: health, the product of health and fractional rank, and the individual's position on the absolute concentration curve, can together be modelled as linear in parameters. As is the case for standard OLS, linearity implies that the marginal effects are constant along the distribution of X and the derivative of Eq. (27) with respect to the covariates X equals the coefficient ψ:

$$\frac{dE[RIF(H, F_Y; \nu^I)|X = x]}{dx} = \frac{d[X'\psi + \mu]}{dx} = \psi, \tag{28}$$

and the unconditional I partial effect equals ψ:

$$\gamma(\nu^I) = \int_{-\infty}^{\infty} \frac{d[X'\psi + \mu]}{dx} \cdot dF(x) = \psi \tag{29}$$

Thus, under the linearity and zero conditional mean assumptions, the marginal effect and the unconditional partial effect are the same and RIF regression is optimally estimated using OLS. The procedure of RIF-I-OLS first involves estimating the empirical RIF, as we outlined in the final part of Section 3. This yields empirical estimates of each individual's recentered influence on I. Then, using the empirical RIF as the dependent variable in an OLS regression we yield the unconditional I partial effects. In practical terms the distinction between the marginal effect and unconditional partial effect becomes important when one relaxes the linearity assumption. As RIF-I-OLS estimates are a first-order approximation of the effect of X on I, the unconditional I partial effect is a local effect estimate of a small change in X. That RIF-I-OLS is a local estimate implies that it should only be considered for relatively small changes. The definition of relatively small will depend on the empirical context, for example the degree to which the true functional form is non-linear and/or the importance of general equilibrium effects.

5. An empirical illustration of WDW decomposition and RIF-I-OLS

In this section we aim to empirically illustrate what the RIF function is, and how WDW decomposition and RIF-I-OLS compare in their interpretation. We also show how RIF-I-OLS is both a well-suited method for determining the causal effect of a covariate on I given a suitable identification strategy and a useful descriptive decomposition method when no causal inference can be made. The illustrative example presented here focuses on the effect of education on income-related health inequality controlling for age and gender and uses data on monozygotic ("identical") twins.

The data is a replica of the data used in Gerdtham et al. (2016). Performing a WDW decomposition, Gerdtham et al. (2016) find education to be significantly associated with a higher level of health and to significantly contribute to the level of inequality, however, this all but disappears when controlling for family and genetic fixed effects common to twin pairs using a twins differencing strategy. To see if these results hold subject to a theoretically less restrictive decomposition method, we extend the analysis by decomposing income-related health inequality using RIF-I-OLS. As in Gerdtham et al. (2016), we first apply a naïve selection on observables identification strategy using OLS and then a twin fixed effects identification strategy.[18] We use the former primarily to illustrate the difference in the interpretation of the results of the two methods but also because most decomposition studies tend to use OLS and even in this descriptive setting RIF-I-OLS has important advantages. We use the twin fixed effects identification strategy to highlight that RIF-I-OLS is well suited to reduced form causal impact analysis where RIF-I-OLS potentially has the most to offer. First, however, we introduce the data and illustrate the empirical RIF of I (focusing on EI in particular).

5.1. Data material

The data used in this empirical example is a subset from the Swedish Twin Registry consisting of respondents that took part in a telephone interview, including a question on self-assessed health, called Screening Across the Lifespan Twin study (SALT) conducted between the years 1998–2002. The final sample size includes 3328 twin pairs born between the years 1931–1958. The survey data is matched with registers from Statistics Sweden on annual taxable gross income (income from earnings, own business, parental leave benefits, unemployment insurance and sickness benefits) and education level. Register data should have relatively small measurement error, which is very important as measurement errors are magnified when differencing between twins, as we do here in the final part of this section. Income is measured as an average of gross income over ages 35–39 years.[19] The education variable is measured as years of schooling and ranges between 8 and 20 years of schooling.[20] To obtain a health measure appropriate for a rank dependent index, we

[18] We refer the reader to Gerdtham et al. (2016) for both an up-to-date discussion on the merits of twin design based studies in revealing the treatment effect of education and for more detailed discussion of the dataset and the twin based fixed effects methodology.

[19] This point is discussed further in Gerdtham et al. (2016).

[20] Years of schooling is imputed from register data using the highest educational degree obtained in the year 1990 as outlined in the appendix in Gerdtham et al. (2016).

Table 1
Variable descriptions, 1st moments and algorithm weights.

Variable	Description	Mean	Algorithm weight
Health	Health utility from TTO algorithm	0.916	
Health1	1 = Very Good Health (self assessed)	0.379	(Reference)
Health2	1 = Good Health (self assessed)	0.37	−0.0315
Health3	1 = Fair Health (self assessed)	0.169	−0.1414
Health4	1 = Poor Health (self assessed)	0.064	−0.3189
Health5	1 = Very Poor Health (self assessed)	0.018	−0.4817
Age4044	1= aged between 40 and 44 years	0.083	0.0109
Age4554	1= aged between 45 and 54 years	0.427	0.0179
Age5564	1= aged between 55 and 64 years	0.449	0.0235
Age6567	1= aged between 65 and 67 years	0.042	0.0193
Female	1 = female, 0 = male	0.551	0.0058
Schooling	Number of years in education	11.571	
Income	Gross income (35–39 years)[a]	199,145	
Constant		1	0.9589

Notes:
[a] Income is in 2010 prices, SEK.

cardinalise the categorical self-rated health measure using a linear algorithm from Burström et al. (2014) (see model 3, supplementary table 8 of their paper) that transforms self-rated health to a time trade-off (TTO) quality of life utility value. The algorithm values taken from Burström et al. (2014) are shown in column 4 of Table 1.[21] Summary statistics are also presented in Table 1.[22]

5.2. *Empirical estimation of the RIF*

The Erreygers index (EI) for estimated health utility scores is 0.03 (Table 2) indicating that higher health utility is more concentrated amongst the rich. The empirical RIF for EI of health utility score ranked by income is calculated as explained in Section 3 and the result is shown in a scatter plot in Fig. 1. Each scatter point in Fig. 1 is an individual's recentered influence value of EI plotted against their income rank. If an individual were to be removed from the sample, the influence on the statistic would be minus that individual's RIF value weighted by the inverse of the sample size. The figure shows that those at the extreme ends of the income distribution have greatest influence on the EI. This is similar to the findings in Monti (1991) for income concentration as measured by the Gini (a univariate rank dependent index): individuals whose income value is at the extremes of the income distribution have greatest influence on the Gini. As the EI is a bivariate index, health, in addition to the ranking variable, affects the degree of influence an individual has on EI. In this particular example those with very poor health (squares) *and* income levels at the extreme ends of the distribution are the ones with the greatest influence on EI.[23] This result is important to note for researchers and policy makers and whilst it may be known to some, the RIF allows it to be shown as a figure. Researchers estimating a rank dependent index as a measure of socioeconomic related health inequality need to be sure that the observations with the largest influence on the statistic are not miss-codings. Policy makers may want to focus attention towards those individuals they can help with most influence on inequality – the extreme poor with poor health in this instance.

5.3. *Interpretation of RIF decomposition and comparison with WDW decomposition*

To provide more information on the characteristics of the individuals that are influencing the statistic, either positively or negatively and to a greater or lesser extent, one may plot the RIF against another variable or turn to the RIF regression method.[24] Table 2 reports descriptive decomposition results of WDW decomposition of EI, and RIF-EI-OLS decomposition, in addition to results for RIF-I-OLS for AC, ARCI, SRCI, and WI alongside standard mean regression. In a descriptive RIF-I-OLS decomposition the estimated coefficients $\hat{\psi}$ may be interpreted as an association between the covariate and the influence on *I*, providing valuable information as to which groups of individuals influence the inequality index. If we (naively) assume the error term, ϵ, and covariates, *X*, are independent having controlled for selection on observables then the RIF-I-OLS parameter $\hat{\psi}$ identifies the (causal) unconditional *I* partial effects of a shift in the distribution of *X* on *I*. Thus, interpretation of $\hat{\psi}$ is similar to the interpretation of the coefficients in standard mean regression (the results of which are shown in column (1) of Table 2). Indeed, RIF decomposition of the mean of health, assuming a health function linear in parameters, is standard OLS (Firpo et al., 2009).

In the decomposition analysis, years of schooling enters the model as an explanatory variable alongside age, gender and interview year dummies (because each twin was not necessarily interviewed at the same time). We only control for age and gender because these variables are exogenous and predetermined before school was attended thereby avoiding the issue of "bad controls" (see 3.2.3 of Angrist and Pischke, 2008). Proceeding in this way allows us to interpret the education coefficient in a meaningful way. Even for descriptive analysis, care should be taken not to introduce mediators that may complicate interpretation. It is for this reason we do not include employment

[21] Health economists often value health states of people by the TTO method where respondents value quality of life in relation to length of life; respondents are asked to imagine living in a given state of health for (typically) ten years, and then to state the shorter amount of time in full health which makes them indifferent between the two options (Drummond et al., 2005). Reference categories are very good self rated health, age 18–24 years and female.

[22] Gerdtham et al. (2016) show that the Swedish Twin Registry data used here is fairly representative of Sweden's population more widely, which otherwise may be a concern for twin based datasets.

[23] The bivariate rank dependent index gives zero weight to those at the median rank (which is the mean ranking value) and increasing weight to those further away from the median. Health values at the mean (in this case those with good health or thereabouts) also have zero impact. This is because the covariance is driven by those furthest away from the mean of the two variables.

[24] One could for example plot a Lowess curve of the RIF and explanatory variable to visually assess a potential relationship and any functional form assumption. We did this for education but there was no real relationship by years of education and therefore do not report the results here.

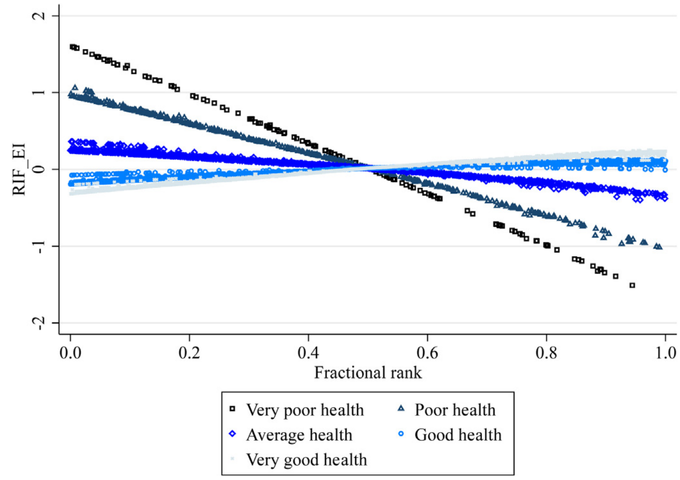

Fig. 1. Scatter plot of individual RIF of EI values plotted against individual's fractional income rank. Each scatter point represents an individual's recentered influence on EI plotted against their fractional income rank by health value.

Table 2
RIF-I-OLS and WDW decomposition estimates of years of schooling, age and gender on income related health inequality.

	OLS	RIF-I-OLS decomposition					WDW EI–OLS decomposition	
Statistic	Health (1)	AC (2)	EI (3)	ARCI (4)	SRCI (5)	WI (6)	Contribution (7)	% contribution (8)
Years schooling	0.005***	0.000	0.001	0.000	0.009***	0.010***	0.008***	0.282***
	(0.000)	(0.000)	(0.001)	(0.000)	(0.003)	(0.003)	(0.001)	(0.035)
Age	−0.000*	0.000	0.001	0.000	0.002	0.003	0.001	0.022
	(0.000)	(0.000)	(0.001)	(0.000)	(0.002)	(0.002)	(0.000)	(0.014)
Male	0.009***	0.003***	0.014***	0.004**	0.051***	0.054***	0.006***	0.212***
	(0.002)	(0.001)	(0.005)	(0.001)	(0.016)	(0.017)	(0.002)	(0.050)
Constant	0.930***	−0.016	−0.066	−0.018	−0.181	−0.199		
	(0.018)	(0.015)	(0.058)	(0.016)	(0.167)	(0.183)		
Mean of RIF	0.916	0.007	0.030	0.008	0.089	0.098		
Observations	6656	6656	6656	6656	6656	6656	6656	6656
WTP FE	NO	NO	NO	NO	NO	NO	NO	NO

Notes: Each column represents a separate decomposition. Column 1 is OLS of the health variable, which is RIF decomposition of the mean assuming linearity in parameters and is optimally estimated using OLS. AC = absolute concentration index, EI = Erreygers Index, ARCI = Attainment relative concentration index, SRCI = Shortfall relative concentration index, WI = Wagstaff Index, WDW = Wagstaff, Van Doorslaer and Watanabe (2003) decomposition. The mean of RIF is the value of the statistic being decomposed. All decompositions control for year of interview fixed effects. Robust standard errors in parenthesis for RIF-mean-OLS and bootstrap standard errors in parenthesis for RIF-I-OLS, 999 repetitions with replacement. Bootstrap standard errors are calculated by bootstrapping the whole procedure (Both for RIF and WDW procedures). Testing null of the coefficient/contributions/% contribution: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

status, for example, as an explanatory variable. Employment status predicts health but it is also an outcome variable affected by education. Its inclusion complicates the interpretation of the education coefficient. The coefficient estimates from RIF-EI-OLS in column (3) of Table 2 suggest, if interpreted as the unconditional I partial effect, that if one made an equal marginal increase to the number of years of education for everyone in the population, this would have no discernible effect on EI. There also appears to be no age profile regarding EI.

Importantly and in contrast to the contribution estimates of WDW decomposition, RIF-I-OLS identifies the effect of the covariates X on the full statistic. That is, the parameter estimate $\hat{\psi}$ captures the effect of the covariates on the *product* of the AC (which is two times the covariance of the level of health and fractional rank) *and* the weighting function $\omega_I(h)$. The parameter estimates $\hat{\psi}$ presented in Columns 2–6 of Table 2 also vary between rank dependent indices depending on the weighting function. Education is found to be significantly associated with the RIF of WI and SRCI, but not with the RIF of AC, EI, or ARCI. That is, more educated individuals have larger influence on the inequality index when measured as WI and SRCI, but not when the AC, EI, or ARCI are considered. This highlights an important issue. The differences in weighting functions, and hence value judgements, among the inequality indices can also lead to important differences in the decomposition results. In this particular example the judgement of whether to consider attainment relative inequality or shortfall relative inequality has bearing on whether education has a potential impact.[25] It is worth noting that it is possible to identify the effect of a particular weighting function by comparing decomposition results for I with decomposition results for AC: AC has a constant weighting function, so any differences in the (standardised) decomposition results compared to those for the AC will be due to the weighting function.

[25] Note that the results divide the indices into two groups. On the one hand EI, AC, ARCI, and on the other hand WI and SRCI. This is a consequence of the high mean of the health utility index.

The last two columns in Table 2 report the results from WDW decomposition of EI. The interpretation is different from any standard form of mean decomposition and to RIF-I-OLS decomposition. The marginal interpretation of WDW ($\upsilon^{\omega_I}(F_H)\beta_k$) (as described in Eq. (10)) implies that a change in the covariance of the covariate and the socioeconomic rank (due to a change in the distribution of the covariate) affects EI by a factor of 4×0.005, with regard to education in this application. It also implies that a change in beta, the health return to the covariate of interest, affects EI by a factor of $4 * 2cov(X, F_Y)$. The procedure also summarises I as a summation of the contribution of each covariate, where these are the covariate-rank covariances weighted by a linear health-covariate correlation. Following the standard practice, we report the WDW decomposition results as contributions from the covariates in levels and percentagewise contributions of the total index. The results suggest that about 28% of the income-related inequalities in health is due to income-related inequalities in education. The contribution is statistically significant suggesting that eliminating income-related inequalities in education might reduce the EI of health, *assuming no change in the ranking variable and a linear health function*.[26] As the procedure ignores the potential impact of the covariates on the weighting function, $\upsilon^{\omega_I}(F_H)$, the percentagewise “contributions” are the same no matter the choice of I (only levels vary with the weighting function). That is, WDW decomposition of any inequality measure implicitly decomposes an absolute index such as EI or AC.

Whilst the results of the two decomposition methods are not directly comparable, as they have different units of measurement, they nevertheless draw contrasting conclusions. WDW decomposition finds a significant contribution due to education whereas RIF for the EI or AC – the most comparable basis, as WDW decomposition holds the weighting function constant and EI and AC have constant weighting functions – finds no significant effect of education. In this particular case we are focussing on a covariate that is well known to causally impact the level of income. Indeed using within twin pair fixed effects on slightly different sample of the same twins population Isacsson (1999) found a significant impact of years of schooling on income and this is a generally accepted finding (Card, 1999). It is hard to interpret the results of WDW decomposition when one knows that a key identifying assumption does not hold (rank ignorability). This is not an uncommon situation; most covariates that impact health also impact the ranking variable. It is our view that the results obtained from RIF regression in this kind of situation are much clearer in their interpretation. RIF-I-OLS results allow us to conclude that there is no *local* association of education with absolute income related health inequality, but there is a *local* association with relative short-falls income related health inequality.

5.4. *The causal effect of education on income-related health inequality*

In the previous section, our identification of the unconditional partial effects did not use twin fixed effects but instead (naïvely) relied on selection on observables to satisfy the assumption that the errors are independent of the covariates. To highlight the importance of causal inference in decomposition analysis we now apply a twins differencing strategy that allows unobserved heterogeneity common between twins to be differenced out. That is, we control for factors such as innate ability and early life factors common to both twins, which may invalidate the exogeneity assumption and yield biased parameter estimates. In the case of income-related health inequality the concern is specifically that this unobserved heterogeneity may be correlated with education and the weighted *covariance* of health and income rank.

To formally derive the within twin pair (WTP) fixed effect decomposition, we denote the RIF values of the jth twin pair, $RIF(h, F_Y(y); I)_{1j}$ and $RIF(h, F_Y(y); I)_{2j}$. Further, we let u_j denote unobserved factors that vary between twin pairs but not within pairs, such as genetic characteristics and certain early life environmental factors and e_{1j} and e_{2j} denote unobserved factors specific to each twin. Assuming a linear functional form for the RIF, we may write these as:

$$RIF(h, F_Y(y); I)_{1j} = X'_{1j}\psi + u_j + e_{1j} \tag{30}$$

$$RIF(h, F_Y(y); I)_{2j} = X'_{2j}\psi + u_j + e_{2j} \tag{31}$$

where X_{1j} is a $k \times n$ matrix of covariates for the first twin in the twin pair j, X_{2j} is for the second twin in the twin pair and ψ is a $k \times 1$ vector of unconditional I partial effects. Taking the difference yields the WTP estimator:

$$RIF(h, F_Y(y); I)_{1j} - RIF(h, F_Y(y); I)_{2j} = \left(X_{1j} - X_{2j}\right)' \psi_{WTP} + e_{1j} - e_{2j} \tag{32}$$

where ψ_{WTP} is the WTP estimator of the effect of education. The unobserved factors that are common to both twins such as genetics or environmental exposure captured by u_j will be differenced out of the equation yielding an unbiased OLS-estimator of $\hat{\psi}$ (given that these are the only sources of unobserved heterogeneity).[27] Applying the WTP approach to the RIF of EI using OLS yields the RIF-EI-WTPFE estimator.

Table 3 reports the monozygotic WTP fixed effects results for EI, AC, CI, and WI alongside standard mean fixed effects regression and WDW decomposition. The results for the RIF-I-WTPFE decomposition suggest that if one gave an equal marginal increase in the number of years of education to everyone in the population, this would have no discernible effect on any measure of I, nor the mean. It therefore appears that either education has no effect on income-related health inequality, or possibly better put: the variation in education that exists under an extensive egalitarian education system cannot explain the observed income-related health inequality.

[26] In the case of EI, the weighting function is a constant and therefore the condition that the weighting function is constant is not binding in this case. However WDW decomposition of CI and WI would also assume that the weighting function is a constant, which it is not.

[27] For a further discussion of potential sources of unobserved heterogeneity see Gerdtham et al. (2016).

Table 3
RIF-I-WTPFE and WDW-WTPFE decomposition estimates of years of schooling on income related health inequality.

Statistic	OLS-WTPFE	RIF-I-WTPFE decomposition					WDW -EI- WTPFE decomposition	
	Health (1)	AC (2)	EI (3)	ARCI (4)	SRCI (5)	WI (6)	Contribution (7)	% contribution (8)
Years schooling	0.001	−0.000	−0.002	−0.000	−0.004	−0.004	0.001	0.05
	(0.001)	(0.001)	(0.002)	(0.001)	(0.009)	(0.007)	(0.002)	(0.058)
Constant	0.930***	0.020	0.081	0.022	0.256	0.277		
	(0.034)	(0.029)	(0.117)	(0.032)	(0.452)	(0.401)		
Mean of RIF	0.916	0.007	0.030	0.008	0.089	0.098		
Observations	6656	6656	6656	6656	6656	6656	6656	6656
WTP FE	YES	YES	YES	YES	YES	YES	YES	YES

Notes: Each column represents a separate decomposition. Column 1 is simply OLS with FE of the health variable. AC = absolute concentration index, EI = Erreygers Index, ARCI = Attainment relative concentration index, SRCI = Shortfall relative concentration index, WI = Wagstaff Index, WDW= Wagstaff, Van Doorslaer and Watanabe (2003) decomposition. The mean of RIF is the value of the statistic being decomposed. All decompositions control for year of interview fixed effects. Robust standard errors in parenthesis for OLS-WTPFE and bootstrap standard errors in parenthesis for RIF-I-WTPFE and WDW, 999 repetitions with replacement. Bootstrap standard errors are calculated by bootstrapping the whole procedure (Both for RIF and WDW procedures). Testing the null of the coefficient/contributions/% contribution: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

6. Discussion

Having introduced and illustrated both the WDW decomposition and RIF-I-OLS decomposition, we now compare the two approaches by summarising the underlying identifying assumptions and differences in interpretation. For clarity, we start by giving a side-by-side comparison of the identifying assumptions of the two approaches.

WDW identifying assumptions:

I. The determinants of health do not determine rank.
II. The determinants of health do not determine the weighting function.
III. Health can be modelled as a function linear in variables X and an error term.
IV. Exogeneity: The errors from the health regression have zero conditional mean.

RIF-I-OLS identifying assumptions:

I. I is differentiable and the differential is bounded.
II. $RIF(h, F_Y(y); I)$ can be modelled as a linear in parameters function of X and an additive error term
III. Exogeneity: The errors from the RIF OLS regression have zero conditional mean.

It is clear from the comparison that RIF-I-OLS requires fewer, and less restrictive, identifying assumptions than WDW decomposition. The first condition for RIF-I-OLS holds as shown in the proof. Exogeneity is of huge importance for causal inference and is common to both methods–but both methods may be used as descriptive exercises without this assumption. Linearity is also common to both methods, but this is an assumption often applied in wider empirical practice and the immediate available evidence suggests this is not a particularly limiting assumption to impose (Van de Poel et al., 2009, Van Doorslaer et al., 2004b, Firpo et al., 2009). This fits with the perceived wisdom that OLS generally provides a good approximation. However, the remaining identifying assumptions of WDW decomposition (rank and weighting function ignorability) are often restrictive as illustrated in this paper and in Erreygers and Kessels (2013). When concern lies with covariates that are known to impact on the ranking variable and the weighting function WDW is likely to yield biased results, which is not true for RIF-I-OLS.

An additional benefit of RIF regression is that it is familiar in its interpretation. The results in Tables 2 and 3 highlight how education's effect on income-related health inequality as estimated by RIF-I-OLS can be shown alongside its effect on mean health in a consistent manner. Similar to mean OLS regression coefficients, the RIF-I-OLS coefficients should be interpreted as how a marginal shift in the distribution of a covariate, e.g., education, influences the inequality index. The interpretation of RIF-OLS estimates and mean OLS regression estimates are similar because RIF-OLS estimates of the mean are in fact exactly the same as mean OLS estimates. We can use this fact to illustrate the difference between WDW decomposition and RIF-I-OLS. WDW is based on a mean OLS regression of health. The contribution of covariate k in WDW decomposition corresponds to its coefficient in an OLS regression on the mean of health - weighted by the weighting function and twice the covariance between covariate k and rank, i.e., $\beta_k \upsilon^{\omega_I}(F_H) 2cov(X_k, F_Y)$. WDW is therefore equivalent to a RIF-OLS decomposition of the mean of health weighted by two functions that are themselves not decomposed. In comparison, RIF-I-OLS estimates the impact of covariates on the index itself, the weighted covariance between health and rank, and therefore decomposes all parts of the index.

As a result of not imposing weighting function ignorability, RIF-I-OLS has the benefit that it allows the analyst to assess the impact of covariates on different forms of I. RIF-I-OLS includes the impact of the covariates on the weighting function and therefore the importance of the covariates may differ between particular indices. Indeed, we illustrate this in our empirical application based on the simple correlations (not WTP fixed effects), where we find that education had no association with the AC, EI and ARCI, but had a significant association with WI and SRCI. Whereas WDW decomposition only allows for decomposing an absolute index, RIF-I-OLS allows researchers to explore how the policy impacts on the level of inequality and how this differs depending on the particular value judgement and hence the particular inequality index policy makers sympathise with. We view this as a necessary part of any inequality analysis because there is no consensus as to which inequality measure is preferred.

In this paper we have highlighted the identifying assumptions of the WDW decomposition method and shown they rarely hold in practice and that this makes interpretation difficult. It appears that the central issue with WDW decomposition is that the parameter of interest is not clear and consequently neither are the conditions under which it will be identified. As Erreygers and Kessels (2013) implicitly show, this results in many potential decomposition methods that can yield very different results. Erreygers and Kessels (2013) conclude with a warning that until it is understood which form of WDW type of decomposition is preferred, all decomposition methods should be used with caution. It is our view that this conclusion should be made a little more explicit and that decomposition methods that *are unclear as to what they estimate and what the necessary identifying assumptions are* should be used with caution. Our approach differs from currently available decomposition methods for bivariate rank dependent indices: first, we are clear as to what our parameter of interest is (the unconditional I partial effect); second, we derive a decomposition method that yields this parameter, based on RIF regression.

On a much simpler level RIF also allows useful graphical presentation of the data. In the empirical example we showed who influence the statistic most – those with very poor health and very low income. Those with very good health have very little impact on the statistic no matter their income rank. This may not be immediately obvious to practitioners and policy makers and is a great way to illustrate who would have greatest impact if targeted. RIF also allows the statistic to be plotted against another covariate and simple bivariate plots can show any potential relation that may be of interest.

The RIF approach does have its limitations; the leading one is that it is a local approximation. It is therefore not reasonable to calculate percentagewise contributions using RIF regression. The usefulness of a local estimate should, however, be placed into the larger context of the overall aims of decomposition analysis and the available alternative approaches. The approach suggested by Kessels and Erreygers (2015) potentially solves the rank ignorability issue but ignores the weighting function assumption. It also requires a structural model, but if one were to have a structural model to hand it may not be preferable to decompose a bivariate index of health inequality. Instead, Fleurbaey and Schokkaert (2009) convincingly make the case for a structural model approach to be used for analysing fair and unfair inequalities in health. As a road map for the health inequality literature this may very well be the goal or ideal that we all should be aiming for. However, if no structural model is available or feasible it may still be of huge interest how a policy change (which is most often a marginal one) impacts both average health and health inequality. RIF of I regression allows this reduced form type of analysis to be made without the need for restrictive assumptions making it a useful addition to a health economist's toolkit.

7. Conclusion

In this paper we have summarised the literature that has identified the identifying assumptions required by WDW decomposition and presented evidence that these assumptions can be important for the decomposition results. Causal analysis using WDW decomposition is therefore troublesome. Even when WDW decomposition is interpreted purely as a descriptive accounting exercise the evidence suggests that results from the WDW decomposition will be difficult to interpret if one is concerned about the rank ignorability and weighting function assumptions. We have introduced an alternative rank dependent index decomposition method that simultaneously relaxes the rank and weighting function ignorability assumptions. This alternative is based on a RIF regression. We have extended the RIF concept from a univariate setting to a general bivariate rank dependent index, providing a method that yields the unconditional I partial effect of a shift in the distribution of X on the inequality index and has strong links to the program evaluation literature. This new decomposition approach is simple to estimate and the interpretation resembles that of standard conditional mean analysis. Our empirical application using the Swedish Twin Registry found a discrepancy between the results of the two methods: WDW decomposition finds a significant association of education and income related health inequality, RIF regression finds no such association. In this example we know that education impacts the ranking variable (income) and therefore interpretation is muddied by the assumptions imposed by WDW decomposition. In comparison, interpreting the results from RIF regression is much clearer and the results suggest there is no *local* impact of education on income related health inequality. In an attempt to illustrate RIF-I-OLS's close link to the treatment effects literature, we used linear WTP fixed effects and found little evidence that (twin differences in) education causally impact income-related health inequality in Sweden.

Finally, it is worth noting that the usefulness of the RIF regression goes beyond the estimation of unconditional I partial effects using OLS. One can for example use instrumental variables for endogenous variables by adding control functions as per Rothe (2010) to obtain consistent estimates of the marginal effects. RIF regression also allows Oaxaca-blinder type decompositions of between group/time differences to be decomposed for statistics other than the mean under some further identifying assumptions (Fortin et al., 2011). We have not discussed these in any great detail but they highlight the potential of our suggested decomposition method and its applicability to a wide range of empirical questions.

Acknowledgments

We are grateful to Peter Lambert for his help expanding on his derivation of the directional derivative for the Gini coefficient. We are also grateful for comments from Sixten Borg, Jens Dietrichson, Andreas Dzemski Guido Erreygers, Roselinde Kessels, Anton Nilsson, Tom Van Ourti, Dennis Petrie, participants at seminars at Lund University, SHEA 2014, University of Duisburg-Essen and University of Antwerp, Essen Health Conference 2015 and two anonymous referees. Financial support from the Swedish Research Council for Working Life and Social Research (FORTE) (dnr 2012-0419), The Swedish Research Council (dnr 2014-646), the Government Grant for Clinical Research ("ALF"), Region Skåne (Gerdtham), and Foundation in memory of Pontus Roos (Kjellsson) is gratefully acknowledged. All remaining errors are the authors' own.

Appendix A. Derivation of the *RIF* for a general rank dependent index (*I*), the IF for the AC and the RIFs for AC, EI, CI, ARCI, SRCI and WI.

Proposition 1. *Let $\nu^I(F_{H,F_Y}) = \nu^{\omega_I}(F_H)\nu^{AC}(F_{H,F_Y})$ be a general rank dependent index, the AC be defined as $\nu^{AC}(F_{H,F_Y}) = 2cov(H, F_Y)$ and F_{H,F_Y} be the joint CDF of H and F_Y with corresponding pdf denoted as f_{H,F_Y}. Then the RIF for $\nu^I(F_{H,F_Y})$ is given by:*

$$RIF(h, F_Y(y); \upsilon^I) = \upsilon^I(F_{H,F_Y}) + IF(h; \upsilon^{\omega_I}) * \upsilon^{AC}(F_{H,F_Y}) + \upsilon^{\omega_I}(F_H) * IF(h, F_Y(y); \upsilon^{AC}),$$

where $IF(h; \upsilon^{\omega_I})$ *denotes the IF of the weighting function for I and* $IF(h, F_Y(y); \upsilon^{AC}) = -2\upsilon^{AC}(F_{H,F_Y}) + \mu_H - h + 2hF_Y(y) - 2\int^y \int^{+\infty} hf_{H,F_Y} dhdF_Y(z)$ *denotes the IF for AC.*

Proof. To show $RIF(h, F_Y(y); \upsilon^I) = \upsilon^I(F_{H,F_Y}) + IF(h; \upsilon^{\omega_I}) * \upsilon^{AC}(F_{H,F_Y}) + \upsilon^{\omega_I}(F_H) * IF(h, F_Y(y); \upsilon^{AC})$, we first apply the definition of the IF given by Eq. (17) to I yielding:

$$IF(h, F_Y(y); \upsilon^I) = \frac{d}{d\varepsilon}\left[\upsilon^{\omega_I}(F_H)\upsilon^{AC}(F_{H,F_Y})\right]\Big|_{\varepsilon=0} \tag{A1}$$

Applying the product rule to Eq. (A1) yields:

$$IF(h, F_Y(y); \upsilon^I) = IF(h; \upsilon^{\omega_I}) * \upsilon^{AC}(F_{H,F_Y}) + \upsilon^{\omega_I}(F_H) * IF(h, F_Y(y); \upsilon^{AC}) \tag{A2}$$

As per Eq. (18), adding the functional $\upsilon^I(F_{H,F_Y})$ to Eq. (A2) yields the RIF for $\upsilon^I(F_{H,F_Y})$.

To show that $IF(h, F_Y(y); \upsilon^{AC}) = -2\upsilon^{AC}(F_{H,F_Y}) + \mu_H - h + 2hF_Y(y) - 2\int^y \int^{+\infty} hf_{H,F_Y} dhdF_Y(z)$, we first note that the absolute concentration index can be written as:

$$\upsilon^{AC}(F_{H,F_Y}) = 2cov(H, F_Y) = 2\int hF_Y dF_{H,F_Y} - 2\int hdF_{H,\infty}\int F_Y dF_{\infty,F_Y} \tag{A3}$$

Eq. (A3) states that AC is a functional of the joint probability distribution F_{H,F_Y} and the probability distribution F_Y. Substituting $\upsilon^{AC}(G_{h,F_y})$ and $\upsilon^{AC}(F_{H,F_Y})$ for $\upsilon^I(G_{h,F_y})$ and $\upsilon^I(F_{H,F_Y})$ in the formula for the bivariate IF given by Eq. (17) yields:

$$IF(h, F_Y(y); \upsilon^{AC}) = lim_{\varepsilon\to 0}\frac{2\int hG_y dG_{h,F_Y(y)} - \int hdG_{h,\infty}\int G_y dG_{\infty,F_Y(y)} - cov(H, F_Y)}{\varepsilon} \tag{A4}$$

Substituting G_y as defined in Eq. (12) and $G_{h,F_Y(y)}$ as defined in Eq. (15) into Eq. (A4) yields:

$$IF(h, F_Y(y); \upsilon^{AC}) = lim_{\varepsilon\to 0}2\frac{\left[\int h((1-\varepsilon)F_Y + \varepsilon\delta_y)d((1-\varepsilon)F_{H,F_Y} + \varepsilon\delta_{h,F_Y(y)}) - \int hd((1-\varepsilon)F_{H,\infty} + \varepsilon\delta_{h,\infty})\int((1-\varepsilon)F_Y + \varepsilon\delta_y)d((1-\varepsilon)F_{\infty,F_Y} + \varepsilon\delta_{\infty,F_Y(y)}) - cov(H, F_Y)\right]}{\varepsilon} \tag{A5}$$

Which after taking the limit and re-arranging yields:

$$\begin{aligned} IF(h, F_Y(y); \upsilon^{AC}) = 2\Bigg[&-2\left(\int hF_Y dF_{H,F_Y} - \int hdF_{H,\infty}\int F_Y dF_{\infty.F_Y}\right) + \int hdF_{H,\infty}\int F_Y dF_{\infty,F_Y} - \int hd\delta_{h,\infty}\int F_Y dF_{\infty,FY} \\ &+ \int hF_Y d\delta_{h,F_Y(y)} - \int hdF_{H,\infty}\int F_Y d\delta_{\infty,F_Y(y)} + \int h\delta_y dF_{H,F_Y} - \int hdF_{H,\infty}\int \delta_y dF_{\infty,F_Y}\Bigg] \end{aligned} \tag{A6}$$

Term by term Eq. (A6) is equal to:

$$-2\int hF_Y dF_{H,F_Y} + 2\int hdF_{H,\infty}\int F_Y dF_{\infty,F_Y} = -\upsilon^{AC}(F_{H,F_Y}), \tag{A7}$$

$$\int hdF_{H,\infty}\int F_Y dF_{\infty,F_Y} = \frac{\mu_H}{2}, \tag{A8}$$

$$-\int hd\delta_{h,\infty}\int F_Y dF_{\infty,F_Y} = -\frac{h}{2}, \tag{A9}$$

$$\int hF_Y d\delta_{h,F_Y(y)} = hF_Y(y), \tag{A10}$$

$$-\int hdF_{H,\infty}\int F_Y d\delta_{\infty,F_Y(y)} = -\mu_H F_Y(y), \tag{A11}$$

$$\begin{aligned} -\int hdF_{H,\infty}\int \delta_y dF_{\infty,F_Y} &= -\mu_H\int^{+\infty}\int^{+\infty}\delta_y f_{\infty,F_Y} dhdF_Y(y) = -\mu_H\int_y^{+\infty}\int^{+\infty} 1f_{\infty,F_Y} dhdF_Y(z) = -\mu_H\int^{+\infty}\int^{+\infty} 1f_{\infty,F_Y} dhdF_Y(y) \\ &+ \mu_H\int^y\int^{+\infty} 1f_{\infty,F_Y} dhdF_Y(z) = -\mu_H + \mu_H F_Y(y), \end{aligned} \tag{A12}$$

$$\begin{aligned} \int h\delta_y dF_{H,F_Y} &= \int^{+\infty}\int^{+\infty} h\delta_y f_{H,F_Y} dhdF_Y(y) = \int_y^{+\infty}\int^{+\infty} h\delta_y f_{H.F_Y} dhdF_Y(z) = \int^{+\infty}\int^{+\infty} hf_{H,F_y} dhdF_Y(y) - \int^y\int^{+\infty} hf_{H,F_Y} dhdF_Y(z) \\ &= -\mu_H - \int^y\int^{+\infty} hf_{H,F_Y} dhdF_Y(z). \end{aligned} \tag{A13}$$

Together these yield:

$$IF(h, F_Y(y); \nu^{AC}) = -2\nu^{AC}(F_{H,F_Y}) + \mu_H - h + 2hF_Y(y) - 2\int^{y}\int^{+\infty} hf_{H,F_Y}dhdF_Y(z), \tag{A14}$$

This completes the proof. □

Corollary 1. *The RIFs for the AC, EI, CI, ARCI, SRCI and the WI are given by:*

$$RIF(h, F_Y(y); \nu^{AC}) = \nu^{AC}(F_{H,F_Y}) + IF(h, F_Y(y); \nu^{AC})$$

$$RIF(h, F_Y(y); \nu^{EI}) = \nu^{EI}(F_{H,F_Y}) + \frac{4}{b_H - a_H} IF(h, F_Y(y); \nu^{AC})$$

$$RIF(h, F_Y(y); \nu^{CI}) = \nu^{CI}(F_{H,F_Y}) + \frac{(\mu_H - h)}{\mu_H^2} * \nu^{AC}(F_{H,F_Y}) + \frac{1}{\mu_H} IF(h, F_Y(y); \nu^{AC})$$

$$RIF(h, F_Y(y); \nu^{ARCI}) = \nu^{ARCI}(F_{H,F_Y}) + \frac{(\mu_H - h)}{(\mu_H - a_H)^2} * \nu^{AC}(F_{H,F_Y}) + \frac{1}{\mu_H - a_H} IF(h, F_Y(y); \nu^{AC})$$

$$RIF(h, F_Y(y); \nu^{SRCI}) = \nu^{ARCI}(F_{H,F_Y}) + \frac{(-\mu_H + h)}{(b_H - \mu_H)^2} * \nu^{AC}(F_{H,F_Y}) + \frac{1}{b_H - \mu_H} IF(h, F_Y(y); \nu^{AC})$$

$$RIF(h, F_Y(y); \nu^{WI}) = \nu^{WI}(F_{H,F_Y}) + \frac{-(b_H - a_H)[(b_H + a_H - 2\mu_H)(h - \mu_H)]}{((b_H - \mu_H)(\mu_H - a_H))^2} * \nu^{AC}(F_{H,F_Y}) + \frac{b_H - a_H}{(b_H - \mu_H)(\mu_H - a_H)} IF(h, F_Y(y); \nu^{AC})$$

Proof. To show the result of Corollary 1, Proposition 1 states the IFs for the weighting functions for the AC, EI, CI, ARCI, SRCI and WI need to be calculated. The weighting functions for both the AC and EI are constants, therefore the IFs for their weighting functions will be zero and we can plug in straight away the functions we need into the formula for the RIF of I. The IF for the CI weighting function is:

$$IF(h; \omega^{CI}) = \frac{d}{d\varepsilon} \frac{1}{\left[(1-\varepsilon)\int hdF_H + \varepsilon h\right]} - \frac{1}{-\int hdF_H}\Big|_{\varepsilon=0}. \tag{A15}$$

Differentiating Eq. (A15) and taking the limit with respect to ε and noting that $\int hdF_H = \mu_H$ gives us:

$$IF(h; \omega^{CI}) = \frac{\int hdF_H - h}{\int hdF_H \int hdF_H} = \frac{(\mu_H - h)}{\mu_H^2} \tag{A16}$$

Substituting Eq. (A16) into the formula for the RIF for I yields the RIF for CI:

$$RIF(h, F_Y(y); \nu^{CI}) = \nu^{CI}(F_{H,F_Y}) + \frac{(\mu_H - h)}{\mu_H^2} * \nu^{AC}(F_{H,F_Y}) + \frac{1}{\mu_H} IF(h, F_Y(y); \nu^{AC}) \tag{A17}$$

The IF for the ARCI weighting function is:

$$IF(h; \omega^{ARCI}) = \frac{d}{d\varepsilon} \frac{1}{\left[(1-\varepsilon)\int (h - a_H)dF_H + \varepsilon(h - a_H)\right]} - \frac{1}{-\int (h - a_H)dF_H}\Big|_{\varepsilon=0} \tag{A18}$$

Differentiating Eq. (A18) and taking the limit with respect to ε gives us:

$$IF(h; \omega^{ARCI}) = \frac{\int (h - a_H)dF_H - (h - a_H)}{\int (h - a_H)dF_H \int (h - a_H)dF_H} = \frac{(\mu_H - h)}{(\mu_H - a_H)^2} \tag{A19}$$

Substituting Eq. (A19) into the formula for the RIF for I yields the RIF for ARCI:

$$RIF(h, F_Y(y); \nu^{ARCI}) = \nu^{ARCI}(F_{H,F_Y}) + \frac{(\mu_H - h)}{(\mu_H - a_H)^2} \nu^{AC}(F_{H,F_Y}) + \frac{1}{\mu_H - a_H} IF(h, F_Y(y); \nu^{AC}) \tag{A20}$$

Following a similar argument as for ARCI, the IF for the SRCI is given by:

$$IF(h; \omega^{SRCI}) = \frac{\int (b_H - h)dF_H - (b_H - h)}{\int (b_H - h)dF_H \int (b_H - h)dF_H} = \frac{(-\mu_H + h)}{(b_H - \mu_H)^2} \tag{A21}$$

Substituting Eq. (A21) into the formula for the RIF for I yields the RIF for SRCI:

$$RIF(h, F_Y(y); \nu^{SRCI}) = \nu^{SRCI}(F_{H,F_Y}) + \frac{(-\mu_H + h)}{(b_H - \mu_H)^2} * \nu^{AC}(F_{H,F_Y}) + \frac{1}{b_H - \mu_H} IF(h, F_Y(y); \nu^{AC}) \tag{A22}$$

The IF for the WI weighting function is given by:

$$IF(h;\omega^{WI}) = \frac{d}{d\varepsilon}\left[\frac{b_H - a_H}{\left(b_H - \int hd\left((1-\varepsilon)F_H + \varepsilon\delta\right)\right)\left(\int hd\left((1-\varepsilon)F_H + \varepsilon\delta\right) - a_H\right)} - \frac{b_H - a_H}{(b_H - \mu_H)(\mu_H - a_H)}\right]|_{\varepsilon=0} \tag{A23}$$

Expanding gives us:

$$IF(h;\omega^{WI}) = \frac{d}{d\varepsilon}\left[\frac{b_H - a_H}{(b_H(1-\varepsilon)\mu_H + b_H\varepsilon h - b_H a_H - (1-\varepsilon)^2\mu_H^2 - (1-\varepsilon)\varepsilon h\mu_H + (1-\varepsilon)a_H\mu_H - (1-\varepsilon)\varepsilon h\mu_H - \varepsilon^2 h^2 + \varepsilon a_H h)}\right.$$
$$\left. - \frac{b_H - a_H}{(b_H - \mu_H)(\mu_H - a_H)}\right]|_{\varepsilon=0} \tag{A24}$$

Differentiating with respect to ε and taking the limit w.r.t. ε yields:

$$IF(h;\omega^{WI}) = \frac{-(b_H - a_H)[(b_H + a_H - 2\mu_H)(h - \mu_H)]}{((b_H - \mu_H)(\mu_H - a_H))^2} \tag{A25}$$

Substituting Eq. (A25) into the formula for the RIF for I yields the RIF for WI:

$$RIF(h, F_Y(y); v^{WI}) = v^{WI}(F_{H,F_Y}) + \frac{-(b_H - a_H)[(b_H + a_H - 2\mu_H)(h - \mu_H)]}{((b_H - \mu_H)(\mu_H - a_H))^2}v^{AC}(F_{H,F_Y}) + \frac{b_H - a_H}{(b_H - \mu_H)(\mu_H - a_H)}IF(h, F_Y(y); v^{AC}), \tag{A26}$$

This completes the proof. □

Appendix B. Linking proposition 1 and corollary 1 to Essama-Nssah and Lambert (2012) and Firpo et al. (2007)

The (R)IF for a univariate rank dependent index, the Gini index (a measure of the concentration of one variable), has been derived in Essama-Nssah and Lambert (2012) and Monti (1991) (only for the IF) and reported in Firpo et al. (2007). If a univariate setting is assumed, where individuals are ranked by health instead of income (i.e., F_H is substituted for F_Y), our derivation of the RIF of the concentration index coincides with previous derivations of the (R)IF of the Gini. As Essama-Nssah and Lambert (2012) show that their result is the same as shown in Firpo et al. (2007), we only need to link our results to the latter.

The IF for the AC is given by proposition 1:

$$IF(h, F_Y(y); v^{AC}) = -2v^{AC}(F_{H,F_Y}) + \mu_H - h + 2hF_Y(y) - 2\int^{y}\int^{+\infty} hf_{H,F_Y}dhdF_Y(z) \tag{B1}$$

If in deriving Eq. (B1) we had used F_H as the ranking variable instead of F_Y we would have got the IF for the absolute Gini index (AG):

$$IF(h, F_H(h); v^{AG}) = -2v^{AG}(F_{H,F_H}) + \mu_H - h + 2hF_H(h) - 2\int^{h}\int^{+\infty} hf_{H,F_H}dhdF_H(z) \tag{B2}$$

Similarly to how the RIF of CI was derived in Appendix A, we find the RIF of the Gini index (GI) equals:

$$RIF(h; v^{GI}) = -\frac{h - 2\mu_H}{\mu_H}v^{GI}(F_{H,F_H}) + \frac{1}{\mu_H}IF(h; v^{GI}) \tag{B3}$$

Rearranging yields

$$RIF(h; v^{GI}) = -\frac{h - 2\mu_H}{\mu_H}v^{GI}(F_{H,F_H}) + \frac{1}{\mu_H}\left[-2v^{AG}(F_{H,F_H}) + \mu_H - h + 2hF_H - 2\int^{h}\int^{+\infty} hf_{H,FH}dhdF_H(z)\right]$$
$$= -\frac{h - 2\mu_H}{\mu_H}v^{GI}(F_{H,F_H}) + -2v^{GI}(F_{H,F_H}) + 1 - \frac{h}{\mu_H} + \frac{2}{\mu_H}hF_H - \frac{2}{\mu_H}\int^{h}\int^{+\infty} hf_{H,F_H}dhdF_H(z) \tag{B4}$$

Note: Firpo et al. (2007) denote the Lorenz ordinate as:

$$\frac{1}{\mu_H}\int^{h}\int^{+\infty} hf_{H,F_H}dhdF_H(z) = \frac{1}{\mu_H}q(\alpha, F_H) \tag{B5}$$

where α is the fractional rank. Firpo et al. (2007) also denote the area under the Lorenz curve as:

$$R(F_H) = \int_0^1 q(\alpha, F_H)d\alpha \tag{B6}$$

The Gini index equals the area between the line of equality and the Lorenz curve:

$$v^{GI}(F_{H,F_H}) = 1 - 2R(F_H) \tag{B7}$$

Substituting Eqs. (B5)–(B7) into Eq. (B5) yields:

$$RIF(h; v^{GI}) = -\frac{(h - 2\mu_H)(1 - 2R(F_H))}{\mu_H} - 2 + 4R(F_H) + 1 - \frac{h}{\mu_H} + \frac{2}{\mu_H}(hF_H - q(\alpha, F_H))] \tag{B8}$$

which after re-arranging yields the expression presented in Firpo et al. (2007):

$$RIF(h; v^{GI}) = 1 + \frac{2hR(F_H)}{\mu_H} - \frac{2}{\mu_H}(h(1 - F_H) + q(\alpha, F_H))] \tag{B9}$$

This completes the proof. □

Appendix C. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jhealeco.2016.03.006.

References

Allanson, P., Petrie, D., 2014. Understanding the vertical equity judgements underpinning health inequality measures. Health Economics 23, 1390–1396.

Angrist, J.D., Pischke, J.S., 2008. Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Burström, K., Sun, S., Gerdtham, U.G., Henriksson, M., Johannesson, M., Levin, L.A., Zethraeus, N., 2014. Swedish experience-based value sets for EQ-5D health states. Quality of Life Research 23, 431–442.

Card, D., 1999. The causal effect of education on earnings. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 3(A), pp. 1801–1863.

Deaton, A.S., 2003. Health, inequality and economic development. Journal of Economic Literature 41, 113–158.

Drummond, M., Sculpher, M., Torrance, G., Brien, O.B., Stoddart, G., 2005. Methods for the Economic Evaluation of Health Care Programmes. Oxford University Press, Oxford.

Erreygers, G., 2009. Correcting the concentration index. Journal of Health Economics 28, 504–515.

Erreygers, G., Kessels, R., 2013. Regression-based decompositions of rank-dependent indicators of socioeconomic inequality of health. In: Rosa Dias, P., O'Donnell, O. (Eds.), Health and Inequality (Research on Economic Inequality), vol. 21. Emerald Group Publishing, London, pp. 227–259 (Chapter 9).

Erreygers, G., Van Ourti, T., 2011. Measuring socioeconomic inequality in health, health care and health financing by means of rank-dependent indices: a recipe for good practice. Journal of Health Economics 30 (4), 685–694.

Essama-Nssah, B., Lambert, P.J., 2012. Influence functions for policy impact analysis. In: Bishop, J. (Ed.), Research on Economic Inequality. , pp. 135–159.

Firpo, S., Fortin, N.M., Lemieux, T., 2007. Decomposing Wage Distributions Using Recentered Influence Function Regressions. Working paper.

Firpo, S., Fortin, N.M., Lemieux, T., 2009. Unconditional quantile regressions. Econometrica 77, 953–973.

Fortin, N.M., Lemieux, T., Firpo, S., 2011. Decomposition methods in economics. In: Ashenfelter, O., Card, D. (Eds.), Handbook of Labor Economics, vol. 4a, pp. 1–102.

Fleurbaey, M., Schokkaert, E., 2009. Unfair inequalities in health and health care. Journal of Health Economics 28, 73–90.

Gerdtham, U.-G., Lundborg, P., Lyttkens, C.H., Nystedt, P., 2016. Do socioeconomic factors really explain income-related inequality in health? Applying twin design to standard decomposition analysis. Scandinavian Journal of Economics 118, 25–48.

Gomez, P.M., Lopez-Nicholas, A., 2005. Socio-economic inequalities in health in Catalonia. Hacienda Pública Española 175, 103–121.

Gravelle, H., 2003. Measuring income related inequality in health: standardisation and the partial concentration index. Health Economics 12, 803–819.

Hampel, F.R., 1974. Influence curve and its role in robust estimation. Journal of the American Statistical Association 69, 383–393.

Hosseinpoor, A.R., Van Doorslaer, E., Speybroeck, N., Naghavi, M., Mohammad, K., Majdzadeh, R., Delavar, B., Jamshidi, H., Vega, J., 2006. Decomposing socioeconomic inequality in infant mortality in Iran. International Journal of Epidemiology 35, 1211–1219.

Isacsson, G., 1999. Estimates of the return to schooling in Sweden from a large sample of twins. Labour Economics 6 (4), 471–489.

Jones, A.M., Nicolás, Á.L., 2006. Allowing for heterogeneity in the decomposition of measures of inequality in health. Journal of Economic Inequality 4 (3), 347–365.

Kessels, R., Erreygers, G., 2015. A unified structural equation modeling approach for the decomposition of rank-dependent indicators of socioeconomic inequality of health. Research report 2015/017. UNU-WIDER – United Nations University, Helsinki, pp. 19, ISBN:978-92-9230-902-2.

Kjellsson, G., Gerdtham, U.-G., 2013a. On correcting the concentration index for binary variables. Journal of Health Economic 32, 659–670.

Kjellsson, G., Gerdtham, U.-G., 2013b. Lost in translation: rethinking the inequality equivalence criteria for bounded health variables. In: Dias, P.R., O'Donnell, O. (Eds.), Health and Inequality (Research on Economic Inequality), vol. 21. Emerald Group Publishing Limited, pp. 3–32.

Kjellsson, G., Gerdtham, U.-G., Petrie, D., 2015. Lies, damned lies and health inequality measurements: Understanding the value judgments. Epidemiology.

Lauridsen, J., Christiansen, T., Gundgaard, J., Häkkinen, U., 2007. Decomposition of health inequality by determinants and dimensions. Health Economics 16, 97–102.

Leu, R.E., Schellhorn, M., 2004. The Evolution of Income-Related Health Inequalities in Switzerland over Time. IZA Discussion Paper Series No 1346.

Mackenbach, J.P., Stirbu, I., Roskam, A.J., Schaap, M.M., Menvielle, G., Leinsalu, M., Kunst, A.E., European Union Working Group on Socioeconomic Inequalities in Health, 2008. Socioeconomic inequalities in health in 22 European countries. New England Journal of Medicine 358, 2468–2481.

Mackenbach, J.P., Kulhanova, I., Menvielle, G., Bopp, M., Borrell, C., Costa, G., Deboosere, P., Esnaola, S., Kalediene, R., Kovacs, K., Leinsalu, M., Martikainen, P., Regidor, E., Rodriguez-sanz, M., Strand, B.H., Hoffmann, R., Eikemo, T.A., Ostergren, O., Lundberg, O., et al., 2015. Trends in inequalities in premature mortality: a study of 3.2 million deaths in 13 European countries. Journal of Epidemiology and Community Health 69, 207–217.

McGrail, K.M., van Doorslaer, E., Ross, N.A., Sanmartin, C., 2009. Income-related health inequalities in Canada and the United States: a decomposition analysis. American Journal of Public Health 99, 1856–1863.

Monti, A.C., 1991. The study of the gini concentration ratio by means of the influence function. Statistica 51, 561–577.

Morasae, E.K., Forouzan, A.S., Majdzadeh, R., Asadi-Lari, M., Noorbala, A.A., Hosseinpoor, A.R., 2012. Understanding determinants of socioeconomic inequality in mental health in Iran's capital, Tehran: a concentration index decomposition approach. International Journal for Equity in Health 11, 18.

Rothe, C., 2010. Identification of unconditional partial effects in nonseparable models. Economics Letters 109, 171–174.

Van de Poel, E., O'Donnell, O., Van Doorslaer, E., 2009. Urbanization and the spread of diseases of affluence in China. Economics & Human Biology 7 (2), 200–216.

Van Doorslaer, E., Masseria, C., Koolman, X., Lafortune, G., Clarke, P., Gerdtham, U.-G., Häkkinen, U., Couffinhal, A., Tubeuf, S., Dourgnon, P., Schellhorn, M., Szende, A., Nigenda, G., Arreola, H., Grasdal, A., Leu, R., Puffer, F., Seidler, E., 2004a. Income-Related Inequality in the Use of Medical Care in 21 OECD Countries. OECD Health Working Paper No 14. Paris.

Van Doorslaer, E., Koolman, X., Jones, A.M., 2004b. Explaining income-related inequalities in doctor utilisation in Europe. Health Economics 13, 629–647.

Van Ourti, T., 2004. Measuring horizontal inequity in Belgian health care using a Gaussian random effects two part count data model. Health Economics 13 (7), 705–724.

Van Ourti, T., van Doorslaer, E., Koolman, X., 2009. The effect of income growth and inequality on health inequality: theory and empirical evidence from the European Panel. Journal of Health Economics 28 (3), 525–539.

Wagstaff, A., Paci, P., Van Doorslaer, E., 1991. On the measurement of inequalities in health. Social Science and Medicine 33, 545–557.

Wagstaff, A., 2005. The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality. Health Economics 14, 429–432.

Wagstaff, A., Van Doorslaer, E., Watanabe, N., 2003. On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in Vietnam. Journal of Econometrics 112, 207–223.

Wilcox, R.R., 2005. Introduction to Robust Estimation and Hypothesis Testing, 2nd ed. Elsevier, Amsterdam.