Estimating the Causal Impact of Macroeconomic Conditions on Income-Related Mortality

Gerdtham, Ulf-Göran; Heckley, Gawain; Lissdaniels, Johannes

Working Paper 2020:22

Department of Economics
School of Economics and Management

Estimating the Causal Impact of Macroeconomic Conditions on Income-Related Mortality

Ulf-G. Gerdtham
Gawain Heckley
Johannes Lissdaniels

November 2020

LUND UNIVERSITY

Estimating the Causal Impact of Macroeconomic Conditions on Income-Related Mortality

Ulf-G. Gerdtham, Gawain Heckley & Johannes Lissdaniels[1]

Abstract

To-date the macroeconomic conditions-mortality literature on income-related inequality in mortality has relied on subgroup analysis, mainly using income as a stratification variable, but this nearly always causes selection bias yielding results that are hard to interpret. To solve this bad control problem, we apply a novel technique based on recentered influence function regression of overall income-related mortality measures, like the commonly used concentration index. We also highlight the importance of: i) measurement of relative versus absolute inequality; ii) measurement of inequality by population-level statistics of inequality (concentration indices) versus subgroup analysis; iii) measurement of short versus long-term income. We illustrate these issues and our suggested solution using detailed individual-level administrative data from Sweden. Our findings show that there overall is a (insignificant) counter-cyclical impact on mortality and its income-related inequality. During a sub-period of pronounced and significant counter-cyclical mortality we find support for accompanying counter-cyclical income-related inequality, but only when using short-term income.

JEL codes: I14, E32

Key words: Mortality; Macroeconomic conditions; Unemployment; Recentered influence function; Inequality; Concentration index.

---

[1] Corresponding author (Lissdaniels): Department of Clinical Sciences - Malmö, Lund University, Center for Economic demography, Lund University, Box 7082, SE-220 07 Lund, Sweden. Email: johannes.lissdaniels@med.lu.se. Gerdtham: Department of Clinical Sciences - Malmö, Lund University; Center for Economic demography, Lund University and Department for Economics, Lund University, Box 7082, SE-220 07 Lund, Sweden, Email: ulf.gerdtham@nek.lu.se. Heckley: Department of Clinical Sciences - Malmö, Lund University; Center for Economic demography, Lund University, Box 7082, SE-220 07 Lund, Sweden, Email: gawain.heckley@med.lu.se. Financial support from the Swedish Research Council (dnr 2019-03553 and dnr 2019-06292) and Krapperupp (KR 2020-0043) is gratefully acknowledged (Heckley). Lissdaniels gratefully acknowledges financial support from Crafoord. The Health Economics Program at Lund University receives core funding from Government Grant for Clinical Research, and Region Skåne (Gerdtham). The administrative data used in this paper comes from the Swedish Interdisciplinary Panel (SIP), administered by the Centre for Economic Demography, Lund University, Sweden. All remaining errors are the authors' own.

1 Introduction

There is a vast literature documenting an effect of changes in macroeconomic conditions on population-level mortality (see e.g. Ruhm (2006) for an early review), yet the joint mortality and socioeconomic status (SES) response to business cycles is still largely unknown. If such heterogeneity exists, this can have important implications for understanding the causes behind SES-related inequality in mortality which is a key issue, particularly in the public health field (CSDH, 2008; Mackenbach, 2012).

Two important problems arise in estimating the impact of the business cycle on SES-related mortality in a subgroup analysis setting. The first problem is that the most common measure of SES, income, generally is a bad control variable (Angrist & Pischke, 2008). This is because income itself depends on the overall climate of the economy and that therefore the composition within subgroups, defined by income, will vary with business cycles. In the analysis of the effect of changes in macroeconomic conditions on income-related mortality, this leads to selection bias and the results are therefore hard, or impossible, to interpret. The second problem is that it may be hard to summarize the overall inequality implications from the subgroup analysis.

In a separate literature on the causes of income-related health inequality, careful consideration has been given to the measurement of overall inequality and the value judgements that different measures rely on. It has however to-date proven difficult to combine credible identification strategies of potential causes with the different measures of inequality. For example, Coveney *et al.* (2020) study the evolution of income-related health inequality before and after the Great Recession in Europe, but as the authors acknowledge they are unable to credibly isolate the impact of the great recession-induced changes in macroeconomic conditions on income-related health inequalities.[2]

In this paper, we contribute to the macroeconomic conditions-mortality literature by drawing upon recent insights from the economic literature on health inequality. We suggest a solution to the bad control problem and also a solution to the problem of summarising the inequality impacts encountered in subgroup analysis. We also shed light on two other key issues that have arisen in this literature, namely the sensitivity of the results to different value judgements underpinning the choice of relative or absolute inequalities, and also the role of the indicator of income, and here we focus particularly on short and long-run income measures.

We discuss these issues, highlight their importance and conclude that the findings from the literature on macroeconomic conditions and mortality rest on fragile grounds. To gain a more solid basis of analysis, we propose the employment of a newly developed decomposition method of inequality indices, RIF-I-OLS, which is based on the Recentered Influence Function (RIF) (Heckley *et al.*, 2016). We show how this method can be applied to investigate the causal impact of changes in macroeconomic conditions on income-related mortality over time.

This RIF approach has several of advantages. First, the method avoids the bad control problem, allowing for causal parameter estimates of the impact of the business cycle on the SES-related mortality. Second, RIF-I-OLS summarizes the overall inequality implications of changes in

[2] Moreover, their method only allows for the decomposition of changes in income-related health inequality across two time-periods, which rules out an investigation of the impact of business cycles on health which is essentially a time series phenomenon requiring longer periods of follow up.

macroeconomic conditions of the entire income distribution into one coherent measure. It thereby overcomes the difficulty of inferring the direction and degree to which income-related mortality is affected by the business cycle encountered in subgroup analysis. Third, this approach allows simultaneous comparison of absolute and relative inequality impacts and therefore allows evaluation of different value judgments in a consistent manner.

We illustrate our points by assessing the impact of the business cycle using administrative data on males aged 20-44 years of age in Sweden during the years 1983-2000. During the period we find a counter-cyclical impact of business cycles on mortality and that this coincided with counter-cyclical increases in short-run income-related mortality. However, changes in macroeconomic conditions did not significantly impact income-related inequality in mortality for all value judgements (indices) considered. Using a long-term definition of SES, long-term income, the impacts are less pronounced compared to inequality measured using short-term income. Importantly, long-term income (and even education) are found to be impacted by changes in macroeconomic conditions, so while the use of long-term income (or education) reduces the selection bias that would be encountered in subgroup analysis where these variables are used to stratify the sample into subgroups, they remain bad controls. This highlights the difficulty of finding a measure of individual's SES that isn't a potential bad control when analysing the inequality impacts of business cycles and further motivates the use of RIF-I-OLS as a tool to assess the impacts of macroeconomic conditions on SES-related mortality.

The paper unfolds as follows. In section 2, we provide a background discussion and review the main literature on macroeconomic conditions and mortality, with a focus on studies investigating the differential impact across different SES groups. In section 3, using the potential outcomes framework, we discuss the bad control problem and the suggested solution. In section 4 we present the data material and the empirical strategy, including details of variable constructions and methods used to estimate the causal impact of business cycles on income-related mortality. In section 5, we first report the results of estimating the impact of business cycles on (male) mortality and position in the income distribution. We then account for the effect of macroeconomic conditions on both these factors simultaneously by decomposing the effect of business cycles on inequality using RIF regressions. Lastly, we extend the analysis utilizing a measure of long-term income. We conclude in section 6.

2 Background

In this section we review the literature investigating SES-related heterogeneity in the business cycle effect on mortality and then follow this with three methodological remarks on the empirical practice typically employed in these studies.

2.1 Previous Literature

The literature based on aggregate regional statistics has most commonly considered the effect of changes in regional unemployment rates on all-cause mortality (Ruhm, 2006), with studies typically showing that mortality is pro-cyclical, i.e. indicating that economic booms (recessions) are bad (good) for health (Ariizumi & Schirle, 2012; Gerdtham & Ruhm, 2006; Neumayer, 2004). Often these studies have been extended considering the main cause-specific mortality rates and results stratified by sex and different age groups. In more recent studies, the literature has been enriched by studies based on individual-level data that allow for analysis of the response in mortality to macroeconomic fluctuations across income groups.

Evidence from individual-level studies on the differential income-related mortality response to business cycles is still scarce. Edwards (2008), using data from the U.S. National Longitudinal Mortality Study for 1979–1998, finds weak evidence that low income groups might be relatively more exposed to procyclical mortality than those with higher income. Conversely, using education as the SES indicator, his results suggest that higher education groups are hit more strongly by pro-cyclical mortality while lower education groups even experience counter-cyclical mortality. Using administrative data from different regions in Norway for 1977–2006, Haaland & Telle (2015) find evidence showing that socioeconomically vulnerable groups have less pro-cyclical mortality than the more advantaged groups. Lastly, van den Berg et al. (2017), using administrative data on Swedish males from 1993 to 2007, find that lower income groups are more affected by pro-cyclical mortality than higher income groups, which they take as an indication that inequality in mortality decreases in recessions and increases in booms.

In sum, it appears that lower income groups are more exposed to pro-cyclical mortality than higher income groups, during periods where pro-cyclical mortality patterns are observed. We continue with three remarks on the above literature and what this implies for its interpretation.

2.2 Remarks on the Previous Literature

1) A challenge in interpreting the results of SES-based subgroup analysis of the mortality response to the business cycle is that income (as typically used) is a bad control. That is because income itself is a function of the "treatment" variable. Hence, its inclusion in a regression of the impact of business cycles on mortality induces selection bias into the model with the unfortunate consequence that the estimated effects do not have a causal interpretation. This is because we do not know who has moved in and out of the subgroups due to changes in macroeconomic conditions, and as a consequence the implications of the selection effect on the results is unknown. It can be shown that even long-term SES indicators, like education or long-term income, are to some extent still affected by macroeconomic activity, especially so for younger adults. We discuss this in more detail and suggest a solution in section 3.

2) Studies that have attempted to estimate the SES-related mortality impacts of business cycles have generally not been explicit about whether the measures under consideration capture absolute or relative effects. One or the other is presented often without justification. For example, Haaland & Telle (2015) find that low SES groups have less pro-cyclical mortality than higher SES groups based on semi-elasticities, i.e. a relative measure, though their results are indeed reversed in absolute terms, that is, based on for example earnings in their Table 5, mortality increases with 82 cases in the low SES group in upturns and 53 cases in the high SES group. In this example, the differences in effect between SES groups were at most borderline significant with the authors concluding the overall pro-cyclicality of mortality was not driven by the people in disadvantaged SES groups.

This example illustrates that the conclusions drawn from an empirical study may well depend on whether absolute or relative inequality measures are used, and relative and absolute measurements also reflect two different normative positions and it is not obvious which one is the most relevant. According to the developing consensus in the health inequality literature (see e.g., (Allanson & Petrie, 2014); Harper et al. (2010); (Kjellsson, Gerdtham, & Petrie, 2015)), it is advisable to report and discuss both absolute and relative measures of inequality.

3) One reason for engaging in subgroup analysis based on SES is because we are interested in the inequality implications of changes in macroeconomic conditions. However, it can be difficult to infer in what direction and to what degree the overall SES-related mortality is affected by business cycles using a subgroup analysis. For example, Haaland & Telle (2015) investigate whether the mortality response differs for different SES groups defined by being below or above the median in earnings, income and wealth. Using this approach, given that the number of cases of death in one group exceeds the number in the other group, the degree to which inequality is affected should depend on where in the distribution these cases of death occur, as cases of deaths further out in the tail of the distribution arguably should have a larger impact on inequality than those close to the median. Accounting for the inequality impact is therefore not merely an accounting exercise on the balance of cases of deaths around the median. Furthermore, and what may be less obvious, even though the number of cases of death in one group may exceed the number in the other group, the direction of the inequality impact (pro-poor versus pro-rich) is determined by where in the distribution the cases of deaths occur.[3] Thus, while subgroup analysis is a valuable tool in uncovering differential effects, inferring the overall inequality response to changes in macroeconomic fluctuations is not trivial, which also makes it hard to compare the inequality effect across studies, countries or over time. Thus, subgroup analysis to infer inequality implications has its limitations.

One solution to this limitation is to measure the overall inequality response directly by using an inequality index measure that allows quantification of the overall impact (direction and degree) across the entire SES distribution. In this context, the dominant way to examine what explains the underlying causes of overall inequality is to decompose the inequality index into a function of its (potential) causes. Decomposition has been exploited by Coveney *et al.* (2020) and by Coveney *et al*. (2016) studying the evolution of income-related health inequalities before and after the Great Recession in Europe and Spain, respectively. However, one serious weakness of this technique is that it exploits national variation over only two time periods, and as a consequence, there is no clear way to isolate exogenous variation in the macroeconomic indicator. Thus it is not possible to causally link the SES-related mortality to macroeconomic conditions under this approach, which is carefully pointed out by Coveney *et al.* (2020) themselves. A further caveat with this type of decomposition analysis is that it only allows for absolute inequality to be decomposed.

3 The Bad Control Problem and A Solution

In this section we set out the selection problem arising from heterogeneity analysis of the business cycle impact on mortality when using income to define subgroups. We do this using a simplified application of the potential outcomes framework. We then introduce RIF regression and extend the potential outcomes framework to show how RIF regression combined with measures of income-related mortality (concentration indices) can solve the selection problem and putting us in a position to conclude about the causal effect of business cycles on mortality over the income distribution.

[3] To be clear, assume for simplicity that the population consists of 11 individuals ordered by income rank, with the first two individuals above median as well as the poorest individual experiencing death. As such, the number of cases of death in the high-status group exceeds the number in the low-status group. Nonetheless, both absolute and relative inequality favours the high-status group, using standard measures of inequality that will be introduced in 3.2; the attainment-relative concentration index equals -0.12 and the absolute concentration index equals -0.02.

3.1 The Bad Control Problem in the Current Context

We are interested in the causal effect of business cycles on mortality. By way of our identification strategy, as outlined in section 4.2.1, this causal effect can be estimated and isolated from other factors impacting mortality. That is, conditional on covariates, those exposed to different macroeconomic conditions are similar on average in every way. We are also interested in how changes in macroeconomic conditions impact income-related mortality. However, if we are not careful, we could end up with results that do not tell us much.

To describe this problem, we turn to the potential outcomes framework and define a binary variable for macroeconomic conditions, $D_i = \{0,1\}$ where 0 is non-recession and 1 is recession. We then define a binary indicator M_i for our outcome of interest, mortality:

$$\begin{matrix} M_{1i} & if\ D_i = 1 \\ M_{0i} & if\ D_i = 0 \end{matrix} \tag{1}$$

where M_{0i} is the potential outcome for an individual who wasn't exposed to a recession and M_{1i} is the potential outcome for an individual who was exposed. Let us also define Y_i as a binary indicator of two income groups where 1 denotes high-income and 0 denotes non-high income:

$$\begin{matrix} Y_{1i} & if\ D_i = 1 \\ Y_{0i} & if\ D_i = 0 \end{matrix} \tag{2}$$

These dichotomisations simplify the illustration of the bad control problem in our context without loss of generality. Under the assumption of conditional independence, D_i is assumed randomly assigned and a comparison of conditional averages by different macroeconomic conditions yields our causal estimates:

$$E[M_i|D_i = 1, X] - E[M_i|D_i = 0, X] = E[M_{1i} - M_{0i}|X] \tag{3}$$

$$E[Y_i|D_i = 1, X] - E[Y_i|D_i = 0, X] = E[Y_{1i} - Y_{0i}|X] \tag{4}$$

Both $E[M_{1i} - M_{0i}|X]$ and $E[Y_{1i} - Y_{0i}|X]$ have a causal interpretation, i.e. they yield the causal effect of business cycles on mortality and income respectively. Now let us consider the interpretation of a difference in mean mortality between non-recession and recession exposed individuals, conditional on being in the high-income group. We can estimate this using subgroup analysis conducted by means of interaction analysis or stratification analysis, the argument is the same. By joint independence of our potential outcomes $\{M_{0i}, M_{1i}, Y_{0i}, Y_{1i}\}$ and treatment, D_i, we get:

$$\begin{gathered} E[M_i|D_i = 1, Y_i = 1, X] - E[M_i|D_i = 0, Y_i = 1, X] \\ = E[M_{1i}|Y_{1i} = 1, X] - E[M_{0i}|Y_{0i} = 1, X] \end{gathered} \tag{5}$$

The Right-Hand Side (RHS) of equation (5) above highlights how including income as a control makes interpretation difficult as:

$$\begin{gathered} E[M_{1i}|Y_{1i} = 1, X] - E[M_{0i}|Y_{0i} = 1, X] \\ = E[M_{1i} - M_{0i}|Y_i = 1, X] + \{E[M_{0i}|Y_{1i} = 1, X] - E[M_{0i}|Y_{0i} = 1, X]\} \end{gathered} \tag{6}$$

The first term of the RHS of equation (6) is the causal effect of recessions on mortality for the high-income group. The second term of the RHS in curly brackets is selection bias, reflecting that recessions change the pool of individuals who belong to the high-income group. This is the well-known bad control problem. We do not know the sign and size of the selection into high SES which means that mortality conditional on SES does not have a causal interpretation. One corollary of the bad control problem is that stratification analysis by income requires an indicator of income not impacted by the business cycle. Yet the current (small) literature has not addressed this issue. This naturally casts doubt on the validity of results from prior studies investigating the income-related heterogeneity in the business cycle effect on mortality.

The bad control problem is a problem of selection, as illustrated in (6), occurring in regression analysis because income used as a stratification variable is itself an outcome of the treatment variable, the business cycle.[4] The analysis of business cycles effects on income-related mortality by way of subgroup analysis is therefore not feasible under common empirical conditions. However, if the outcome of interest is population-level inequality and not differences in individual outcomes across subgroups, we can reframe the question and avoid the selection problem. Income-related mortality can be measured in the form of a population statistic and in doing so we move income to the LHS of the regression model to form part of the dependent variable. In doing so we can, in fact, study the co-movements of mortality and income with changes in the business cycle, rather than seeing movements in income as a source of bias and an empirical nuisance, induced by conditioning on SES.

To understand this idea, which is the key contribution of this paper, we introduce how a population-level statistic can be linked to individual-level characteristics using the RIF. We first introduce population-level bivariate inequality indices (section 3.2) and how they can be linked to individual characteristics (section 3.3). We then revisit the potential outcomes framework introduced in this subsection to show how the selection problem is addressed when using RIF combined with bivariate statistics (section 3.4).

3.2 The Concentration Index and its Many Transformations

The dominant way of measuring income-related inequality in the field of health economics is to use the Concentration Index (*CI*) of health, which belongs to a family of bivariate rank dependent indices. These measures are bivariate because an individual's health is related to her income, and they are rank dependent because individuals are sorted according to their rank in the income distribution. The *CI* quantifies the extent to which health overall is systematically related to income by summarizing the relationship between cumulative health and income-rank into one coherent measure, where a positive (negative) index value represents a pro-rich (pro-poor) distribution of health (Wagstaff, Paci, & van Doorslaer, 1991).

A complication in our empirical set-up is that our health indicator, denoted H, is bounded from below by a_H and above by b_H. Let μ_H denoted the mean of H and F_Y the fractional rank of income as an indicator of SES, where F_Y is uniformly distributed over the unit interval with mean ½, then the general form of a rank dependent index (I) is given by:

$$I = \mathcal{W}^I(H) ACI(H) \tag{5}$$

[4] The fact that the bad control variable not only is included as a control variable in the model but is interacted with the business cycle proxy using interaction analysis, or used to stratify the sample using stratification analysis, further complicates interpretation.

where $\mathcal{W}^I(H)$ is a weighting function specific to a particular form of rank dependent index I, and the Absolute CI ($ACI(H)$) is given by twice the covariance between H and F_Y:

$$ACI(H) = 2Cov(H, F_Y) \tag{6}$$

For bounded health variables such as mortality the choice of health outcome, mortality or its opposite - survival, yields different values of relative inequality measurements (Clarke *et al.*, 2002). As a response to this Erreygers (2009) suggested an alternative index, which is an absolute measure of inequality.[5] We follow Kjellsson (2015) and also consider two relative measures of inequality that acknowledge the bounded health variable issue. We therefore choose to use the mortality Attainment-Relative CI ($MRCI$), the mortality Shortfall-Relative CI ($SRCI$) and the Erreygers CI (ECI) with the corresponding weighting functions given by:

$$\mathcal{W}^{EI} = \frac{4}{b_H - a_H} \tag{7}$$

$$\mathcal{W}^{ARCI(H)} = \frac{1}{(\mu_H - a_H)} \tag{8}$$

$$\mathcal{W}^{SRCI(H)} = \frac{1}{(b_H - \mu_H)} \tag{9}$$

Both the $ARCI$ and $SRCI$ are relative measures of inequality and are invariant to equi-proportional changes in health attainments and health shortfalls respectively. $ARCI$ and $SRCI$ have the same signs but different magnitudes as the former is normalized with the mortality rate while the latter is normalized with the survival rate. The added value of considering both $ARCI$ and $SRCI$ is that we consider our measure of health (mortality) whilst simultaneously also considering measurement in terms of attainment or shortfall on the relative measure of health inequality. In contrast, ECI is an absolute measure of inequality as it is invariant to the addition or subtraction of an equal amount of health for all individuals in the population. Unlike $ARCI$ and $SRCI$, ECI is symmetric in attainment and shortfall.

It is important to note that any choice of an index represents a particular value judgement defined by the weighting function $\mathcal{W}^I$(Allanson & Petrie, 2014; Kjellsson *et al.*, 2015). The fact that there does not exist an actual consensus as to which value judgment that should be

[5] Erreygers interprets the aforementioned issue and the fact that the bounds of the index value is contingent of the mean of the health indicator as technical inconsistencies which motivate a correction, hence he defined the Erreygers *CI* . Instead, Kjellsson, Gerdtham, and Petrie (2015) argue that with bounded health variables these properties of the relative *CI* reflect two different kinds of relative value judgements. It is a general result that a relative index can never yield the same level of inequality in attainments and shortfall (exempting the case where $\mu_M = \mu_S$). Thus, for an index to yield the same level of inequality in attainments and shortfall one needs to impose an absolute value judgement of inequality, i.e. defining the weighting function such that its magnitude is the same in both attainment and shortfall.

preferred motivates our choice of presenting two kinds of relative and one absolute measure of income-related mortality.

3.3 RIF Regression of the Various Concentration Indices

In this section, the concept of RIF regression is introduced which allows us to link individual characteristics to the *CI*. The concept of RIF builds upon the concept of the Influence Function (IF) that comes from the robustness statistics literature, allowing analysis of how robust a statistic is to outliers. The IF yields each individual's influence on a statistic, or how the statistic will change resulting from small perturbations of the distribution on which the statistic is defined. More simply, it tells us how the statistic changes due to the removal of individual observations. The RIF is a transformation of the IF, adding the value of the population statistic to the IF, so it is "recentered", rendering the expectation of the RIF vector to be equal to the original statistic of interest.

Let $\nu_{H,Y}$ be functional, a *CI* in our case, with *Y* representing a continuous income measure. The IF of $\nu_{H,Y}$ is $IF(\nu_{H,Y})$ and the RIF is given by:

$$RIF(\nu_{H,Y}) = IF(\nu_{H,Y}) + \nu_{H,Y} \tag{10}$$

Firpo *et al.* (2009) show that because we can apply the law of iterated expectations (LIE) to this vector of recentered influences we can link individual characteristics to any population statistic that we can calculate a RIF for, and this is performed by way of RIF regression. RIF regression coefficients provide an approximation for the effect of a change in the distribution of the covariate on our distributional statistic of interest, $\nu_{H.Y}$. Heckley *et al.* (2016) show that RIF regression can be applied to the *CI* under certain conditions. By the LIE we have:

$$\nu_{H,Y} = E\left[RIF(\nu_{H,Y})\right] = E\left[E\left[RIF(\nu_{H,Y})\middle|X = x\right]\right] \tag{11}$$

For a general function of covariates X and an error term ϵ the conditional expectation of the RIF may then be modelled as:

$$E\left[RIF(\nu_{H,Y})\middle|X = x\right] = \lambda(X, \epsilon) \tag{12}$$

This is RIF regression. We assume additive linearity in covariates and zero conditional mean of the error term and estimate equation (12) by OLS:

$$E\left[RIF(\nu_{H,Y})\middle|X = x\right] = X'\beta \tag{13}$$

where β then yields the marginal effect of X on $\nu_{H,Y}$, that is:

$$\frac{dE\left[RIF(\nu_{H,Y})\middle|X = x\right]}{dx} = \frac{d[X'\beta]}{dx} = \beta \tag{14}$$

And by linearity the unconditional partial effect also equals:

$$\int_{-\infty}^{\infty} \frac{d[X'\beta]}{dx} . dF(x) = \beta \tag{15}$$

The Influence Function is defined for a very small perturbation in the joint distribution $F_{H,Y}$, and by extension the marginal effect and the unconditional partial effect are both defined for a very small perturbation in the distribution of covariates, from F_X in the direction of G_X. In our illustrative example we consider changes that are larger: the difference in the level of health inequality caused by non-recession and recession exposed individuals. We can describe our covariates as two distributions, non-recessionary (F_X) and recessionary (G_X). Given this, the impact of recessions on the CI of mortality consists of the marginal effect of the variable of interest, $\beta(v_{H,Y})$, and a remainder term $\mathrm{r}(v_{H,Y}, G, F)$:

$$\delta(v_{H,Y}) = v_{H,Y}(G) - v_{H,Y}(F) = \beta(v_{H,Y}) + \mathrm{r}(v_{H,Y}, G, F) \tag{16}$$

That is, the true effect of business cycles on our statistic, $\delta(v_{H,Y})$, can be approximated by a linear function of the marginal effect and a remainder term. For small changes the remainder term will be very small. The marginal effect can therefore be considered as the first term of a linear approximation of the true effect. That is the marginal effect provides an approximation of the effect on $\nu_{H,Y}$ had we made a slight change in the distribution of the covariate *x*.

3.4 RIF Regression, Concentration Indices and the Potential Outcomes Framework

In the standard case, discussed in section 3.1, where subgroup analysis is conducted to estimate the impact of business cycles on income-related mortality, we used the potential outcomes framework to consider individual potential outcomes and then extrapolated this to the population as we believe we can say something about treatment effects based on differences across group averages. For the *CI* (a population-level statistic) we do not have individual outcomes, but we do have RIF values, that in expectation equal the original statistic. In this case for any individual we have the two following potential RIF values for the *CI*:

$$\begin{array}{ll} RIF(\nu_{H,Y})_{1i} & if\ D_i = 1 \\ RIF(\nu_{H,Y})_{0i} & if\ D_i = 0 \end{array} \tag{17}$$

Where $RIF(\nu_{H,Y})_{1i}$ is the RIF value of the *CI* for an individual who was exposed to a recession and $RIF(\nu_{H,Y})_{0i}$ is the RIF value of statistic $\nu_{H,Y}$ for an individual who wasn't exposed to a recession. What we observe is:

$$RIF(\nu_{H,Y})_i = RIF(\nu_{H,Y})_{0i} + \left(RIF(\nu_{H,Y})_{1i} - RIF(\nu_{H,Y})_{0i}\right) D_i \tag{18}$$

Similarly, as for the mortality case, under the assumption of conditional independence, we assume that D_i is as good as randomly assigned and a comparison of conditional averages by different macroeconomic conditions yields our causal estimate:

$$\begin{gathered} E\left[RIF(\nu_{H,Y})_i \middle| D_i = 1, X\right] - E\left[RIF(\nu_{H,Y})_i \middle| D_i = 0, X\right] = \\ E\left[RIF(\nu_{H,Y})_{1i} - RIF(\nu_{H,Y})_{0i} \middle| X\right] \end{gathered} \tag{19}$$

Thus, regressing the RIF vector on the business cycle given conditional independence yields the causal marginal effect of the business cycle on the *CI* of mortality. There is no selection problem because income is no longer included in the RHS of the regression where its dependence on the business cycle induces selection bias due to conditioning. This is a simplified example that without loss of generality highlights how the standard bad control problem is solved using the suggested RIF regression approach. In practice the treatment and stratification variables may be non-binary as in the current context, but the implications are the same.

The cost is that the RIF regression coefficient is an *approximation* of how the *CI* changes due to a small shift in the distribution of macroeconomic conditions as described by equation (16). Thus our causal estimates of the impact of changes in macroeconomic conditions on income related mortality inequality given by equation (19) can be considered as a linear first-order approximation of the true change that occurs.

3.5 A Credible Identification of the Impact of Business Cycles on Inequality in Mortality

RIF regression of the *CI* (and its variants) solves the bad control problem yielding credibly identified and interpretable results of the business cycles impact on income-related mortality that have a causal interpretation. This is the major contribution of our approach. Also, it allows us to address the other remarks highlighted in the discussion above. That is because RIF regression can be applied to any form of *CI*, the analysis can be applied to both absolute and relative measures. RIF regression also provides a single figure estimates of the impact of business cycle on income-related mortality, as the *CI* is a population-level statistic. This makes summarizing the impact on *overall* inequality a simpler task compared to subgroup analysis.

4 Data and Empirical Strategy

The data sources and variable construction are first presented in this section. Subsequently, the identification strategies used for estimating the impact of the business cycles on mortality and income-related inequality respectively are presented.

4.1 Data Sources

We use population-based administrative data drawn from the Swedish Interdisciplinary Panel (SIP) that includes the universe of males in the age group 20-44 years resident in Sweden during the period 1979-2000. We focus on young males since this subgroup is expected to be particularly affected by the change in labour market conditions based on previous studies. The dataset is created by merging several registers through personal identifiers, covering information on date and residence of death, demographic variables, income and educational attainment. In addition to the individual-level data from SIP, regional-level data on unemployment rates was collected from Statistics Sweden.

In measuring income, we use annual disposable after-tax family income.[6] As annual income varies on a year-to-year basis, this measure of income can be designated as short-term income. However, since people who die lose the capacity to generate income, their annual income is mechanically lower due to a period of zero income after death in that year. Similar to some previous studies we solve this by using a one-year lag of income. When income is used in subgroup analysis, subgroups are identified using quintiles of the income distribution. To

[6] Equivalized disposable income was not available for the whole sample period

measure various *CI*s, on the other hand, individuals are ranked according to their income. Let Y and R denote income and a function that ranks individual $i = 1, \ldots, n$ in time t, respectively, then the income rank used to calculate the inequality index (I) in year t, is:

$$rank_{it} = R_{it}(Y_{it-1}) \tag{20}$$

Thus, for each year, individuals receive a rank according to their previous years income and this rank is used to calculate yearly measurements of the income-related inequality in mortality.

4.2 Estimation Strategy

In this subsection, we present our empirical estimation strategy to identify the impact of business cycles on mortality as well as its impact on income-related mortality.

4.2.1 Identifying the Impact of Business Cycles on Mortality

The standard model (Ruhm, 2000) for estimating the causal effect of business cycles on mortality takes the following form:

$$H_{ijt} = \alpha + \tau + \mathcal{T} + \beta Unem_{jt} + \varepsilon_{ijt} \tag{21}$$

where subscripts i, j, and t indicate individual, region and year. τ is a vector of year fixed-effects and can be thought of as controlling for nonlinear time-trends in holding constant universal determinants of mortality occurring yearly across regions. α is a vector of regional fixed-effects that controls for time-invariant unobserved heterogeneity that differs across regions. $Unem_{jt}$ is the unemployment rate in region j in year t, and serves as our measure of the business cycle in line with the literature. The impact of business cycles on mortality, β, is therefore identified by within-regional variations, relative to business cycles occurring in other regions. Lastly, $\mathcal{T}$ denotes regional specific times-trends that are included to account for differences in trends in mortality rates between regions that may correlate with changing conditions in regional labour markets. To answer the question of whether different people in different income groups are affected differently by business cycles, subgroup analysis is employed as per estimating (21) on the population stratified by income quintiles.

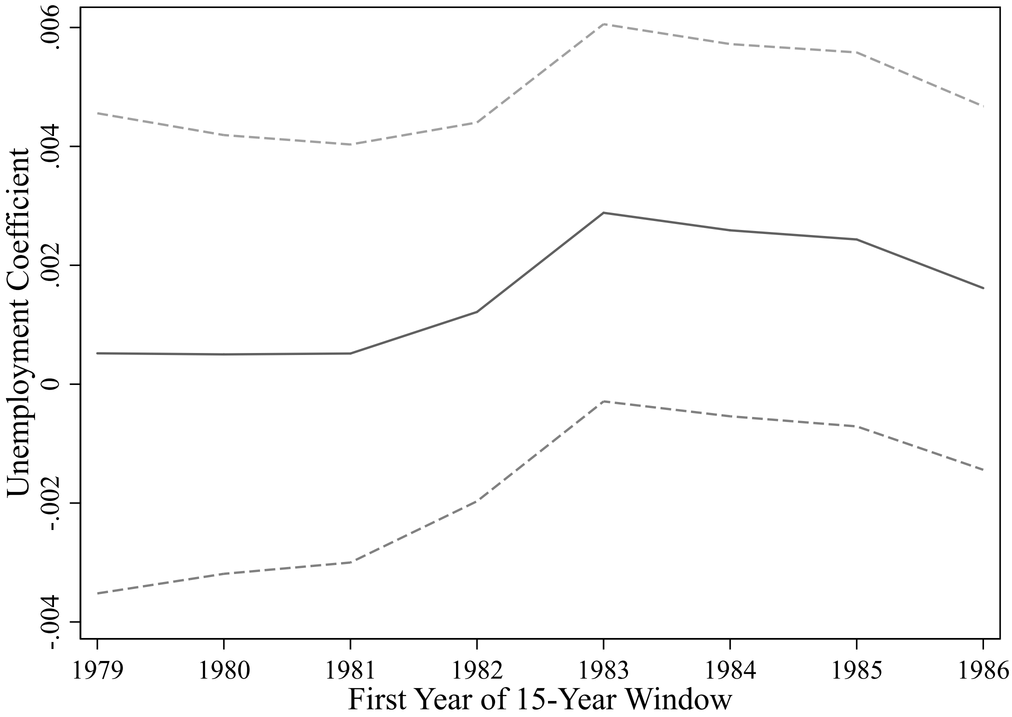


Figure 1. Rolling Regressions of Mortality on Business Cycles on a 15-year Window

Notes: This figure plots the business cycle impact on mortality from an OLS regression of equation (21) for a moving 15-year window, with start year given by the x-axis. Dashed lines represent the robust 95% confidence intervals. Source: Swedish Interdisciplinary Panel and Statistics Sweden, authors' calculation.

Our strategy starts by running an initial regression on mortality and regional unemployment rates for the entire period controlling for regional-specific and year-specific fixed effects as well as regional-specific time trends. The results indicate, in opposite to most of the literature, including a previous study from Sweden based on similar data but from a more recent period (van den Berg *et al.*, 2017), that mortality is weakly countercyclical (the estimate of the unemployment rate is small and significant only at the 10% level). To see if the initial results hold up to scrutiny, we re-estimate the relationship by rolling regressions in eight 15-year data set windows, see Figure 1. Although the estimates vary over time, from window-to-window, our initial finding of countercyclicality holds for all the windows.[7] To help illustrate the methodological issues addressed in the current paper we drop the first four years of the period (1979-1982) and focus on the period 1983-2000 where the estimate of interest is twice the initial estimate and is significant at the 5% level. By focusing on a period when the cyclical mortality pattern is more pronounced, we can more easily illustrate the methodological issues that can arise when estimating income-related impacts of business cycles on mortality. Our empirical results are therefore only valid for a particular period when counter-cyclical mortality was significant. That the impact of business cycle fluctuations on mortality may be sensitive to the time period of analysis has been well documented (see e.g. Ruhm, 2015)

[7] Although indicated, the shift in mortality response in Sweden needs to be corroborated, preferably in one coherent investigation covering at least several decades, an endeavor which may be undertaken using aggregated regional data having the merits of allowing for longer time series not accessible using individual-level data. As the counter-cyclical pattern is consistent during our observation window, it is not possible to further investigate this suggested shift in cyclical pattern from counter-cyclical as encountered in this study to pro-cyclicality in later periods observed in other studies.

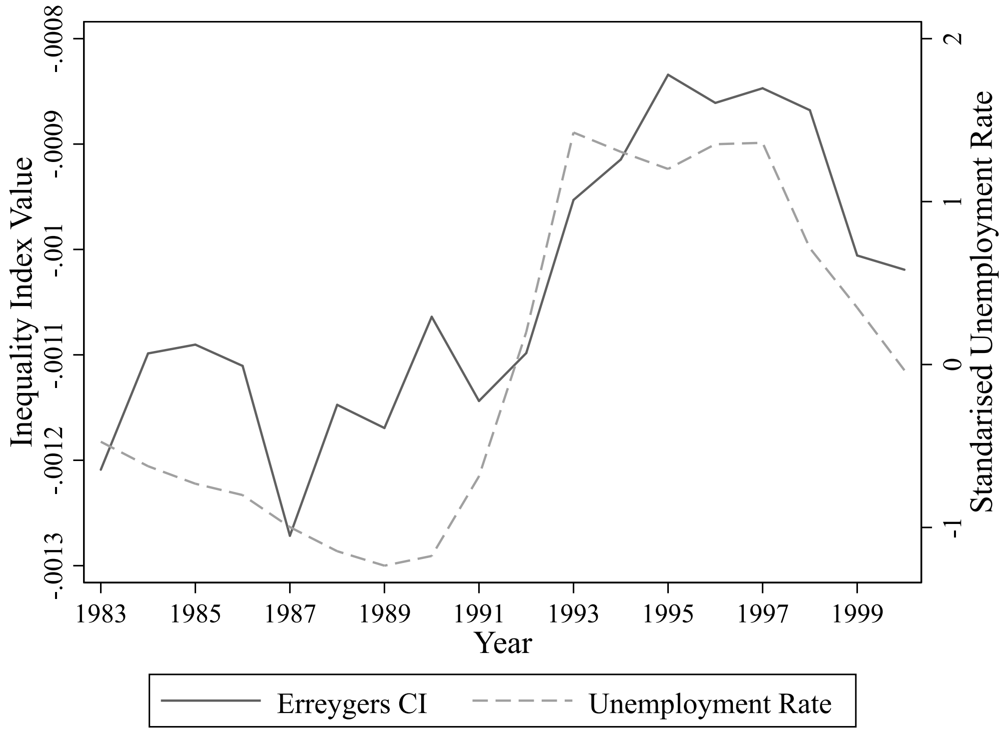
4.2.2 Identifying the Impact of Business Cycles on Inequality in Mortality

Previewing our econometric results, in Figure 2 we plot our measures of income-related mortality for the ECI, $ARCI$ and *SRCI* alongside standardized unemployment rates, all measured for each year of our analysis sample. The average index values for the period are -0.001, -0.22, and -0.0002 for the ECI, *ARCI* and *SRCI* respectively. That is, there was a pro-poor concentration of mortality over the analysis period. During the sample period, the inequality indices decreased in a cyclical fashion. The figures indicate a counter cyclical correlation between unemployment rates and income-related mortality.

To identify the causal link of business cycles on inequality, the RIF is estimated for each CI measure $I = \{ECI, ARCI, SRCI\}$ following Heckley *et al.*, (2016). We choose to analyse inequality at the population level because it is the standard level of analysis for inequality, meaning that we calculate the I for the country as a whole and for each year. To get exogenous variation in the business cycle we exploit regional-level variation as above, rather than variation at the country level that is perfect colinear with all other determinates of mortality and income occurring yearly across the country. Using the RIF vector as a dependent variable in linear regression we are then able to estimate the marginal effect of macroeconomic conditions on the inequality in mortality:

$$RIF(\nu_{H,Y})_{it} = \alpha + \tau + \mathcal{T} + \gamma Unem_{jt} + \epsilon_{ijt}, \tag{22}$$

where $\alpha, \tau, \mathcal{T}$ are defined as per equation (21). The impact of business cycles on our inequality measures is captured by β. Because *ECI*, *ARCI*,and *SRCI* are negative-valued (pro-poor) in our case, the sign of γ has to be considered carefully. A negative γ will mean that an increase in unemployment *increases* both absolute and relative income-related inequality in mortality.


Inequality Index Value
-.0013
-.0012
-.0011
-.001
-.0009
-.0008
Standarised Unemployment Rate
-1
0
1
2
1983
1985
1987
1989
1991
1993
1995
1997
1999
Year
Erreygers CI
Unemployment Rate

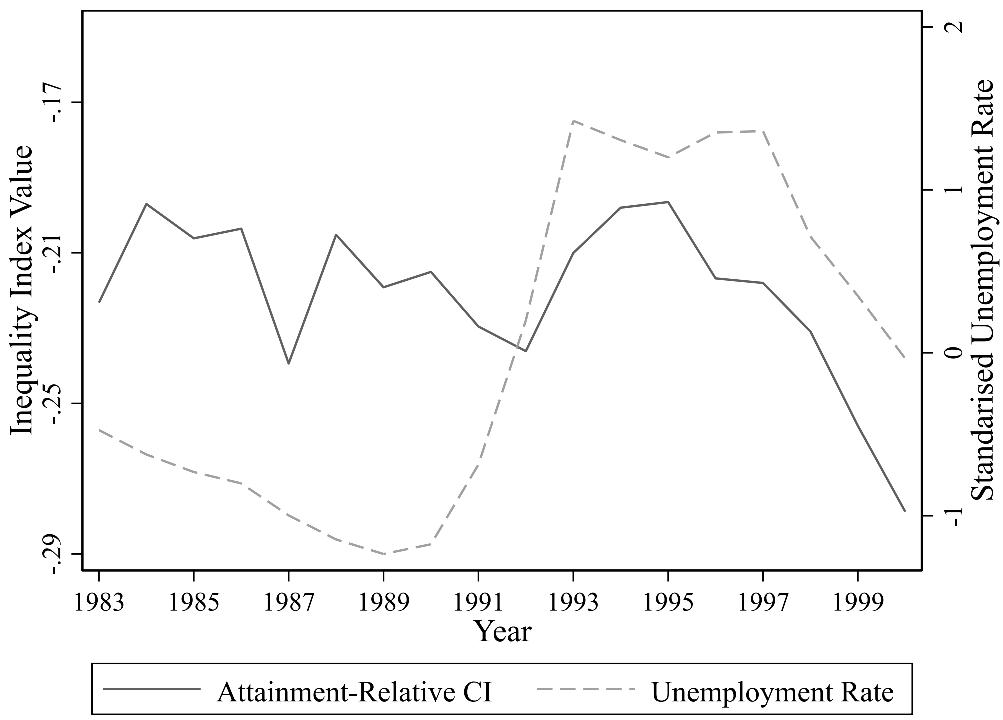

Inequality Index Value
-.29
-.25
-.21
-.17
Standarised Unemployment Rate
-1
0
1
2
1983
1985
1987
1989
1991
1993
1995
1997
1999
Year
Attainment-Relative CI
Unemployment Rate

Figure 2: The Evolution of Income-Related Inequality in Mortality and Unemployment
Notes: This figure shows the co-movement between standardised collapsed regional-level unemployment rates and Erreygers Concentration Index (ECI), the Attainment-Relative (ARCI) and the Shortfall-Relative (SRCI) Concentration Index.

5 Results

The impact of the business cycle on income-related mortality depends on its impact on both mortality and income. We, therefore, start by studying the impact of business cycles on mortality and income separately. Following this, we account for both factors simultaneously by decomposing the effect of business cycles on inequality using RIF regression. Extending this analysis, we explore a measure of long-term income that ought to be less sensitive to changes in business cycles in an attempt to provide credible evidence of the differential mortality response across different income groups. Lastly, the business cycle impact on inequalities in mortality is analysed using this long-term income measure.

5.1 The Causal Impact of the Business Cycle on Mortality

Model specification (1) in Table 1 controls for year fixed-effects only and indicates that mortality increases in recessions, implying a counter-cyclical co-movement of mortality with the business cycle. A one percentage point increase in regional unemployment rates increases the predicted number of deaths with 20 cases, equivalent to a semi-elasticity of 1.1. As we successively add regional fixed-effects (2) and regional linear time-trends (3), the business cycle effect on mortality increases. In specification (3) a one percentage point increase in regional unemployment rates increases the predicted number of deaths to 49 cases, equivalent to semi-elasticity of 2.7 per cent. From now and onwards all results are estimated using our most general model, model (3), where regional time-trends are included in addition to year and region fixed-effects.

Table 1: The Effect of Regional Unemployment Rates on Male Mortality

	(1)	(2)	(3)
Regional unemployment rates	0.00129**	0.00289**	0.00325**
	(0.00052)	(0.00117)	(0.00142)
Constant	0.00112***	0.00104***	0.00102***
	(0.00003)	(0.00006)	(0.00007)
Predicted number of deaths	20	44	49
Semi-elasticity	1,1	2,4	2,7
Year fixed effects	Yes	Yes	Yes
Regional fixed effects	No	Yes	Yes
Regional time trends	No	No	Yes
Observations	27,381,389	27,381,389	27,381,389

Notes: This table presents the effects of regional unemployment rates on male mortality for the period 1983 to 2000. Semi-elasticities are calculated by dividing the marginal effect by the mortality rate (0.0012). Robust standard errors are shown in parentheses.*** p<0.01, ** p<0.05 and * p<0.1 denote significance level.

5.2 The Causal Impact of the Business Cycle on the Short-term Income

To study the impact of changes in macroeconomic conditions on short-term income, a set of linear probability models are estimated. Table 2 shows that the business cycle impact on the probability of belonging to a particular income quintile group.

Table 2. The Effect of Regional Unemployment Rates on Short-Term Income.

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Regional unemployment rates	-0.0409**	-0.0826***	0.238***	-0.115***	-0.00001
	(0.0164)	(0.0166)	(0.0166)	(0.0166)	(0.0164)
Constant	0.202***	0.204***	0.188***	0.206***	0.200***
	(0.00082)	(0.00083)	(0.00083)	(0.00083)	(0.00082)
Observations	27,381,389	27,381,389	27,381,389	27,381,389	27,381,389

Notes: This table presents the effect of regional unemployment rates on the probability of belonging to a short-term income quintile for the period 1983 to 2000, estimated with a linear probability model. Short-term income is defined as the one year lagged disposable family income. The models include controls for the year and regional fixed-effects, and regional time-trends. Robust standard errors are shown in parentheses. *** p<0.01, ** p<0.05 and * p<0.1 denote significance level.

The parameter estimates in table 2 show clear impacts on the quintiles of income. This shows that changes in macroeconomic conditions cause important income rank mobility even for income that is lagged by a year. Indeed, because income exhibits strong serial correlation, using additional lagged values of income would not solve the selection bias issue.

The results in Table 2 illustrate that the use of income quintiles for subgroup analysis of the impact of macroeconomic conditions on mortality will likely lead to selection bias. The information presented in this table is however not enough to tell us what this selection bias may look like. This is because we do not know from where in the income distribution individuals move from or too, nor do we know the respective mortality rates of these individuals. Therefore, the discussion of the bad control problem in section 3 is indeed important and pertinent in the context of macroeconomic conditions and mortality, and that it is highly questionable whether conducting simple subgroup analysis based on these income groups is meaningful. For reference nonetheless, we include these analyses in the web appendix.

5.3 The Impact of the Business Cycle on Inequality in Mortality

The first two columns in Table 3 show the impact of the business cycle on absolute and attainment relative inequality in mortality, measured by *ECI* and *ARCI* respectively. The third column shows the impact on short-fall relative inequality in mortality measured by the *SRCI.* The estimated effects are negative indicating that both absolute and relative inequality increases in recessions, or if you will, decreases during good economic times. With the *ECI* taking the mean value of -.001, a one percentage point increase in unemployment rates is associated with a 5 per cent increase in absolute inequality, while attainment-relative inequality with a mean ARCI of -0.22 would increase by 2 per cent. Turning to inequality in shortfalls, the estimated effect also indicates that relative inequality decreases during recessions, this time with about 7 per cent for a one percentage point increase in unemployment. Evidence for a significant business cycle impact on inequality is found for the *ECI* and *SRCI* at 10-per cent level of significance. To note, the impact of the business cycle on inequality in mortality ceases to be significant when estimated for all years 1979-2000 (results shown in web-appendix) mirroring the findings for mortality.

Table 3. The Impact of the Business Cycle on Short-term Income-Related Inequality in Mortality.

	ECI	ARCI	SRCI
Regional unemployment rates	-0.00541*	-0.401	-0.00136*
	(0.00325)	(0.71471)	(0.000815)
Constant	-0.00077***	-0.201***	-0.00019***
	(0.00016)	(0.0325)	(0.00004)
Observations	27,381,389	27,381,389	27,381,389

Notes: This table presents the effect of regional unemployment rates on male short-term income-related inequality in mortality using the Recentered Influence Function (RIF) of the Erreygers Concentration Index (ECI), the Attainment-Relative (ARCI) and the Shortfall-Relative (SRCI) Concentration Index, as the dependent variable, respectively, in an OLS-regression for the period 1983 to 2000. The RIFs are calculated yearly at the national level using lagged disposable family income as the ranking variable. The models include controls for the year and regional fixed-effects, and regional time-trends. Bootstrap standard errors are shown in parenthesis. The procedure starts by generating a new dataset (999) by random sampling with replacement on which the RIF vectors are calculated that subsequently are regressed on to the business cycle. The standard error of the vector of estimated coefficients constitutes our bootstrap standard errors. *** $p<0.01$, ** $p<0.05$ and * $p<0.1$ denote significance level.

5.4 An Extended Analysis Based on Long-Term Income

A potential way to deal with the bad control problem could be to use long-term income indicators which ought to be less sensitive to temporary changes in macroeconomic conditions which we will explore in this section. We measure long-term income by age and year-standardized individual mean over time of lagged income. The construction of this measure is presented in detail in appendix A.

5.4.1 Subgroup Analysis of Business Cycles on Mortality of Long-Term Income

The results of the subgroup analysis are displayed in Table 4. Following the recommendation presented in the background, we report both absolute and relative measures and, for completeness, we also report group-specific mortality rates. Significant counter-cyclical mortality is indicated in the second quintile, experiencing a rather large mortality response both in absolute and relative terms. A one percentage point increase in regional unemployment rates is associated with 23 additional cases of death in this group with a semi-elasticity of 6.4.

Table 4. The Effect of Regional Unemployment Rates on Mortality by Quintiles of Long-Term Income.

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Regional unemployment rates	0.00350	0.00755**	-0.00165	0.00368	0.00319
	(0.00461)	(0.00319)	(0.00292)	(0.00259)	(0.00218)
Constant	0.00226***	0.00080***	0.00112***	0.00059***	0.00035***
	(0.00023)	(0.00016)	(0.00015)	(0.00013)	(0.00011)
Predicted number of deaths	11	23	-5	11	10
Semi-elasticity	1,4	6,4	-1,6	4,7	6,3
Mortality rate	0,0024	0,0012	0,0010	0,0008	0,0005
Observations	5,406,871	5,475,757	5,494,382	5,502,368	5,502,011

Notes: This table presents the subgroup analysis results of the effects of regional unemployment rates on male mortality where the population is stratified by long-term income quintiles for the period 1983 to 2000. Long-term income is defined as the individual mean of the age and year standardized lagged disposable family income. Semi-elasticities are calculated by dividing the marginal effects by the mortality rates, respectively. The models include controls for the year and regional fixed-effects, and regional time-trends. Robust standard errors are shown in parentheses.*** p<0.01, ** p<0.05 and * p<0.1 denote significance level.

Taken together the estimates and semi-elasticities across all five quintile groups indicate that absolute inequality might increase in recessions while relative inequality might decrease, although it is hard to draw definitive conclusions about the overall impact on inequality.

5.5 The Causal Impact of the Business Cycle on the Long-Term Income

The validity of the results presented in Table 4 rests on the assumption that the composition of the long-term income quintile groups are fixed with respect to macroeconomic activity, or at least that the mobility across quintile groups is unrelated to mortality. We test the assumption of whether the long-term income quintile groups are fixed and provide the results in Table 5, showing that the probability of belonging to the highest two quintiles are affected by changes in unemployment rates. Compared to the relationship between the business cycle and short-term income this relationship appears less pronounced. However, it is worth bearing in mind that there may still be mobility between long-term income groups due to changes in macroeconomic conditions that on average cancels out. The finding of selection into long-term income quintiles due to the business cycle means that we still conclude that the results from the subgroup analysis presented in Table 4, are uncertain since they are hard to interpret.

Table 5 The Effect of Regional Unemployment Rates on Long-Term Income.

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Regional unemployment rates	0.0232	-0.00285	0.0119	-0.127***	0.0942***
	(0.0163)	(0.0166)	(0.0166)	(0.0166)	(0.0163)
Constant	0.196***	0.200***	0.200***	0.207***	0.196***
	(0.00081)	(0.00083)	(0.00083)	(0.00083)	(0.00082)
Observations	27,381,389	27,381,389	27,381,389	27,381,389	27,381,389

Notes: This table presents the effect of regional unemployment rates on the probability of belonging to a long-term income quintile for the period 1983 to 2000, estimated with a linear probability model. Long-term income is defined as the individual mean of the age and year standardized lagged disposable family income. The models include controls for the year and regional fixed-effects, and regional time-trends. Robust standard errors are shown in parentheses.*** p<0.01, ** p<0.05 and * p<0.1 denote significance level.

5.6 The Impact of the Business Cycle on Long-Term Income-related Inequality in Mortality

In this subsection, we assess the impact of business cycles on inequality in mortality based on long-term income. Table 6 presents the results. The sign of the estimates indicates that absolute inequality, and shortfall-relative inequality, increase in recessions, while attainment-relative inequality decreases.[8] Judging by significance, the analyses suggest that long-term based inequality does not vary with the business cycle irrespective of the choice of index measure.

Table 6. The Impact of the Business Cycle on Long-term Income-Related Inequality in Mortality.

	ECI	ARCI	SRCI
Regional unemployment rates	-0.00252	0.641	-0.00063
	(0.00362)	(0.69427)	(0.00090)
Constant	-0.00133***	-0.339***	-0.00033***
	(0.00018)	(0.0347)	(0.00004)
Observations	27,381,389	27,381,389	27,381,389

Notes: This table presents the effect of regional unemployment rates on male long-term income-related inequality in mortality using the Recentered Influence Function (RIF) of the Erreygers Concentration Index (ECI), the Attainment-Relative (ARCI) and the Shortfall-Relative (SRCI) Concentration Index, as the dependent variable, respectively, in an OLS-regression for the period 1983 to 2000. The RIFs are calculated yearly at the national level using the individual mean of the age and year standardized lagged disposable family income as the ranking variable. The models include controls for the year and regional fixed-effects, and regional time-trends. Bootstrap standard errors are shown in parenthesis. The procedure starts by generating a new data set (999) by random sampling with replacement on which the RIF vectors are calculated that subsequently are regressed on to the business cycle. The standard error of the vector of estimated coefficients constitutes our bootstrap standard errors. *** p<0.01, ** p<0.05 and * p<0.1 denote significance level.

[8] To understand the positive ARCI coefficient, note that business cycle impact on the covariance between mortality and rank decreases in the long-term analysis; compare the ECI coefficients in the short and long-term. This means that the relative influence of the weighting function on the marginal effect of ARCI increases, which after weighting by the ACI yields a positive influence in this case, explaining the change to a positive sign. Thus, while there still is a negative correlation between mortality and rank, the magnitude of the pro-poor attainment-relative inequality decreases as the covariance is scaled by a slightly larger morality rate in recessions. This example illustrates the strength of RIF-I-OLS of being able to decompose the marginal effect on different value judgments.

6 Discussion

In this paper, we have argued and also shown, that analysis of the business cycle impact on income-related inequality in mortality by use of subgroup analysis is not advisable, since the results that follow do not have a clear interpretation. This is because income is itself generally a function of the macroeconomic climate and is, therefore, a bad control variable when used to define subgroups, leading to selection bias.

6.1 The RIF Regression Approach to Analyse the Impact of Business Cycles on Income-Related Inequality

To move the business cycle-mortality and income related health inequality literatures forward, we have proposed the analysis of concentration indices where the joint impact of the business cycle on both mortality and income are exploited. By moving the bad control variable out of the conditional expectation of the regression we overcome the problem of selection. This is accomplished by employing RIF regression of the *CI* of income-related mortality. Indeed, this approach allows the assessment of the business cycle impact on income-related mortality, including measures of income that are highly influenced by changes in regional labour market conditions, like current income. Compared to prior descriptive decomposition analyses on income-related health inequality (e.g. Coveney *et al.,* 2020) our method allows for identification of the *causal effect* of changes in macroeconomic conditions on income-related mortality. Using this strategy, we find some support for the conclusion that income-related mortality varies with fluctuations over the business cycle during a period of significant variation of mortality and the business cycle. Over the full sample period with less pronounced variation of mortality and the business cycle, we found no statistically significant variation of income related mortality and the business cycle.

Additional merits of using RIF-I-OLS compared to the standard routine of inferring the inequality implications indirectly through subgroup analysis is that the population-level inequality impact is estimated directly providing us with a single overall measure of the effect; an effect that can easily be compared across different model specifications, over time and across countries and across studies. For example, our results indicate that a one percentage point increase in regional unemployment rates causes an approximately 5 per cent increase in absolute inequality as measured by the ECI. A further advantage is that RIF-I-OLS can be applied to both absolute and relative measures of income-related inequality in a consistent manner, as it in contrast to existing decomposition techniques (e.g. Coveney *et al.,* 2020) can be applied to any *CI*. Indeed, we find statistically significant evidence that absolute inequality and shortfall-relative inequality increases in recessions, that the size of the changes in inequality depend on the measure of inequality, yet no significant effect is found for attainment-relative inequality.

Whilst the RIF-I-OLS approach is advantageous in a number of important ways, the results remain approximations of the true marginal effect of business cycles on the various *CIs*. RIF-I-OLS yields a linear approximation of the marginal effect of business cycles on inequality and is therefore only valid for small changes. Another limitation is that the RIF-I-OLS approach is not able to answer the question of whether those from a lower income quintile are affected more (or less) than those from a higher income quintile. It only informs about the overall inequality impact of the business cycle, which whilst an important question of policy interest, there is also a clear interest in understanding impacts for subgroups. This leads us to discuss other ways of avoiding the selection bias problem in the context of analysing the impact of business cycles on income-related mortality.

6.2 In Search of Measures of SES not Impacted by Business Cycles

We have provided an approach that allows analysis of the business cycle impact on SES-related mortality even for measures of SES that are themselves impacted by macroeconomic conditions. However, an alternative and potentially simpler approach to avoiding the bad control problem is to consider the use of indicators of SES that are either pre-determined with respect to the business cycle or that measure permanent life-cycle income and therefore not influenced by such variations. We discuss several alternatives, their practicality and the consequences of their use on the conclusions that can be drawn.

Using SES prior to the period of the analysis is an apparent solution as it is pre-determined. SES defined before the window of analysis most likely ensures that it is not predicted by changes in macroeconomic conditions, occurring after its measurement. However, given that most business cycle analysis is over a long period (at least 15 years) the relevance of a SES variable set at the beginning of the period for an individual's actual SES many years later is questionable if the results are then going to be translated into policy recommendations. This approach also leads to an important measurement problem, in that younger individuals will not have a measure of SES before the sample window because education, income and occupation are simply not known for children and young adults. Another approach could be to use parental SES prior to the period of analysis instead of own SES, but this modifies the question by shifting focus from own SES to SES background, and this may not be the right way to go, i.e. modifying the analytical question in response to empirical problems. Own SES for older cohorts prior to the period of analysis may however be a more relevant measure. For example, it has been shown that income of middle-aged individuals, in the age range around 35-50 years, are a good predictor of permanent income in Sweden (Böhlmark & Lindquist, 2006**.** The downside with using a measure of permanent income is of course that the validity of this measure depends on the empirical context, meaning that a measure that has shown valid in one setting, for example, one country or subgroup, not necessarily can be assumed to be valid in other settings. For example, using the income of middle-aged individuals in Sweden as a measure for permanent income does not hold for females, only for males.

Instead of using some pre-determined measure of SES or a proxy for permanent income ala (Böhlmark & Lindquist, 2006), we could employ a measure of long-term income, highest occupational status or highest educational level achieved to capture SES. The idea is that a long-term measure of SES is not impacted by the business cycle, but this is an empirical question which can be tested. Indeed, in our own attempt to create a long-term income measure we still found it was predicted by regional unemployment rates. This is theoretically possible because long-term income will be impacted by individual periods of unemployment, the longer-term scarring effects of unemployment and human capital accumulation effects and so on. In line with this, we find that own education is also impacted by the regional unemployment rates (see web appendix), indicating direct human capital accumulation effects of changes in regional labour markets. Furthermore, even if long-term SES is not found to be associated with changes in economic conditions on average, it will still suffer a measurement problem for those who die. Young people who die as a consequence of the business cycle may not have yet completed their educational investments, will have only just started their careers and therefore their long-term income will largely be unknown. The bad control problem is hence not easily solved by using alternative SES indicators and this may even modify the analytical question.

Whilst it is hard to find a measure of SES not impacted by the business cycle, even when considering long-term measures, the choice of SES shouldn't be made primarily based on the empirical challenges at hand. Short-term measures and long-term measures of SES capture different things. SES can be thought of as a relatively stable individual marker that doesn't vary much over time. On the other hand, current SES (e.g. income, employment status) is also of interest, even if temporary, as it can have notable interactions with health. For example, a previously affluent person may have temporarily lost his/her job and subsequently become depressed. This negative association between current income and health is interesting even though the individual's long-term income is of high SES nature. Thus, both short-term and long-term income measures are both of interest but for different questions.

To summarise, care needs to be made in the choice of SES indicator when analysing the impact of the business cycle on SES-related mortality. A clear justification is needed for the choice of SES indicator or preferably, multiple SES indicators should be considered. If subgroup analysis is to be performed, then tests should be used to show that the indicator of SES is not impacted by the business cycle. If the SES indicator is impacted by the business cycle, we currently cannot perform subgroup analysis by SES as the results will not be easily interpretable, but we can still analyse the impact on the overall inequality using RIF-I-OLS.

7 Conclusion

The analysis of the impact of macroeconomic conditions on income-related mortality is thwarted with empirical challenges that without careful consideration could lead to misleading conclusions. To avoid these pitfalls in future research and to help ensure full consideration is given to the issues we have highlighted we suggest that future analysis of macroeconomic conditions and income-related mortality should: 1) when conducting subgroup discuss and test the predicative relationship between the business cycle and income (SES); 2) summarise the impacts on overall inequality using RIF-I-OLS. A complete analysis should also consider both absolute and relative inequalities as well as both short-term and long-term measures of income (SES).

In this paper, we have argued that when analysing the impact of business cycle fluctuations on income-related mortality short-term income is likely to be of great interest but so is a more long-term or permanent measure of income. We have shown that both short-term and even long-term income can be influenced by the business cycle and this makes subgroup analysis based on different income groups difficult due to the problem of selection bias. In this case, one can still provide credible answers to the business cycle impact on overall SES-related mortality using RIF-I-OLS, that in additions allows for estimating impacts on both absolute or relative inequality.

Whilst we move the literature forward by proposing a method that allows us to credibly investigate the impact of the business cycle on income-related mortality we have been unable to provide a method for the analysis of its impact on different income groups when income itself is impacted by the business cycle. We leave this as an interesting problem for future research to potentially solve.

References

Allanson, P., & Petrie, D. (2014). Understanding the vertical equity judgements underpinning health inequality measures. *Health Economics, 23*(11), 1390-1396. doi:10.1002/hec.2984

Angrist, J. and S. Pischke (2008). *Mostly Harmless Econometrics: An Empiricists' Companion.* Princeton, NJ., Princeton University Press.

Ariizumi, H., & Schirle, T. (2012). Are recessions really good for your health? Evidence from Canada. *Social Science & Medicine, 74*(8), 1224-1231. doi:10.1016/j.socscimed.2011.12.038

Böhlmark, *A., & Lin*dquist, M., J. . (2006). Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden. *Journal of Labor Economics, 24*(4), 879. doi:10.1086/506489

CSDH (2008). *Closing the gap in a generation: health equity through action on the social determinants of health. Final Report of the Commission on Social Determinants of Health. Geneva,* World Health Organization.

Coveney, M., García-Gómez, P., Van Doorslaer, E., & Van Ourti, T. (2016). Health Disparities by Income in Spain Before and After the Economic Crisis. *Health Economics, 25 Suppl 2*, 141-158. doi:10.1002/hec.3375

Coveney, M., García-Gómez, P., van Doorslaer, E., & Van Ourti, T. (2020). Thank goodness for stickiness: Unravelling the evolution of income-related health inequalities before and after the Great Recession in Europe. *Journal of Health Economics, 70*. doi:10.1016/j.jhealeco.2019.102259

Clarke, PM, Gerdtham, U, Johannesson, M, Bingefors, K & Smith, L 2002. On the measurement of relative and absolute income-related health inequality. Social Science and Medicine, vol. 55, no. 11, pp. 1923-1928. https://doi.org/10.1016/S0277-9536(01)00321-5

Edwards, R. (2008). Who is hurt by procyclical mortality? *Soc Sci Med, 67*(12), 2051-2058. doi:10.1016/j.socscimed.2008.09.032

Erreygers, G. (2009). Correcting the Concentration Index. *Journal of Health Economics* **28**(2): 504-515.

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

Gerdtham, U., & Ruhm, C. J. (2006). Deaths rise in good economic times: evidence from the OECD. *Economics and Human Biology*(3), 298. doi:10.1016/j.ehb.2006.04.001

Haaland, V. F., & Telle, K. (2015). Pro-cyclical mortality across socioeconomic groups and health status. *J Health Econ, 39*, 248-258. doi:10.1016/j.jhealeco.2014.08.005

Harper, S., King, N. B., Meersman, S. C., Reichman, M. E., Breen, N., & Lynch, J. (2010). Implicit value judgments in the measurement of health inequalities. *Milbank Q, 88*(1), 4-29. doi:10.1111/j.1468-0009.2010.00587.x

Heckley, G., Gerdtham, U. G., & Kjellsson, G. (2016). A general method for decomposing the causes of socioeconomic inequality in health. *J Health Econ, 48*, 89-106. doi:10.1016/j.jhealeco.2016.03.006

Kjellsson, G., Gerdtham, U., & Petrie, D. (2015). Lies, Damned Lies, and Health Inequality Measurements: Understanding the Value Judgments. *Epidemiology*(5), 673. doi:10.1097/EDE.0000000000000319

Mackenbach, J. P. (2012). The persistence of health inequalities in modern welfare states: The explanation of a paradox. *Social Science & Medicine* 75(4): 761-769.

Neumayer, E. (2004). Recessions lower (some) mortality rates:: evidence from Germany. *Social Science & Medicine, 58*(6), 1037-1047. doi:10.1016/S0277-9536(03)00276-4
Ruhm, C. J. (2000). Are Recessions Good for Your Health? *The Quarterly Journal of Economics, 115*(2), 617.
Ruhm, C. J. (2006). Macroeconomic Conditions, Health and Mortality. In A. M. Jones (Ed.), *Elgar Companion to Health Economics* (pp. 5-16.). Cheltenham, UK: Edward Elgar Publishing.
Ruhm, C. J. (2015). Recessions, healthy no more? *J Health Econ, 42*, 17-28. doi:10.1016/j.jhealeco.2015.03.004
van den Berg, G. J., Gerdtham, U. G., von Hinke, S., Lindeboom, M., Lissdaniels, J., Sundquist, J., & Sundquist, K. (2017). Mortality and the business cycle: Evidence from individual and aggregated data. *J Health Econ, 56*, 61-70. doi:10.1016/j.jhealeco.2017.09.005
Wagstaff, A., Paci, P., & van Doorslaer, E. (1991). On the measurement of inequalities in health. *Social Science & Medicine, 33*(5), 545-557. doi:https://doi.org/10.1016/0277-9536(91)90212-U

Appendix A: A measure of long-term SES

The measure of long-term SES position is based on mean income during the observation period. One problem that arises when comparing SES positions of individuals of different ages is however that productivity and income vary over the lifecycle. As a solution to the problem that an individual's rank may be sensitive to their age during the observed period, we age-standardized income analogously to the indirect standardization of a health variable (O'Donnell et al. 2008). This method allows individuals to have their own age but the same mean age effect on income as the entire population. This is accomplished by estimating the following income regression:

$$Y_{it-1} = b_1 age_{it} + b_2 age_{it}^2 + b_3 age_{it}^3 + b_4 age_{it}^4 + \varepsilon_{it} \tag{23}$$

where the age polynomial is of a sufficiently high order to capture non-linearities in the age that we want to standardize for. The indirect standardization is given by the difference between the actual and the age-expected income plus the overall sample mean:

$$Y_{it-1}^{IS} = Y_{it-1} - \bar{Y}_{it-1}^{age} + \bar{Y}_{it-1} \tag{24}$$

Age adjustment is the first step. In the second step, we account for the fact that real income follows economic growth with Sweden being a richer county at the end of the observed period and therefore incomes are not immediately comparable across time. This problem is solved by simultaneously standardizing for year effects in (23). One may note that with this strategy variation in income caused by national business cycle fluctuations is absorbed in the year effects. As we employ a regional fixed effects estimator as per (21) this is not a problem since the estimator only exploits changes in regional business cycles (relative to other regions). Our final long-term income is therefore given by:

$$\bar{Y}_i^{IS} = \frac{1}{\sum_i \tau_i} \sum_{t=1983}^{2000} Y_{i,t-1}^{IS} \tag{25}$$

where $\bar{Y}_i^{IS}$ is the age and sample year standardized mean of previous year income for each individual, τ_i[9] is an indicator variable for the years an individual is in the sample. Hence, $\bar{Y}_i^{IS}$ measures an individual's average or long-term income. Finally, in order to receive yearly ranks used in the calculation of I, we apply the ranking function to $\bar{Y}_i^{IS}$,

$$\tilde{R}_{it} = R_{it}(\bar{Y}_i^{IS}) \tag{26}$$

where $\tilde{R}_{it}$ is the rank of individual i in year t of $\bar{Y}_i^{IS}$. The reason why $\tilde{R}_{it}$ changes over time for individual i while $\bar{Y}_i^{IS}$ does not is because the cross-sectional cohorts where relative SE position is measured change from year to year as new individuals enter the analysis, those turning 19 years old, while older individuals exit the analysis, those turning 40 years old.

[9] For example, for an individual that is 20 years old in the year 1990 and who is alive the year 2000, $\sum_i \tau_i$ equals 11.

8 Web Appendix

Table A1 The Effect of Regional Unemployment Rates on Educational Attainment.

	14 years	15.5 years	19 years	21 years
Regional unemployment rates	-0.144***	0.0652***	0.00222	0.00576**
	(0.0184)	(0.0142)	(0.00191)	(0.00282)
Constant	0.161***	0.0929***	0.00166***	0.00407***
	(0.00120)	(0.000934)	(0.000125)	(0.000185)
Observations	16,399,055	16,399,055	16,399,055	16,399,055

Notes: This table presents the effect of regional unemployment rates on the probability of year of education for the period 1990 to 2000, estimated with a linear probability model. The models include controls for the year and regional fixed-effects, and regional time-trends. Robust standard errors are shown in parentheses.*** p<0.01, ** p<0.05 and * p<0.1 denote significance level.

Table A2. The Effect of Regional Unemployment Rates on Mortality by Quintiles of Short-Term Income.

	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Regional unemployment rates	0.00701*	0.000338	0.00826***	0.00103	-0.00118
	(0.00393)	(0.00374)	(0.00298)	(0.00259)	(0.00233)
Constant	0.00143***	0.00166***	0.000635***	0.000737***	0.000683***
	(0.000194)	(0.000189)	(0.000151)	(0.000129)	(0.000117)
Additional number of deaths	21	1	25	3	-4
Semi-elasticity	3,9	0,2	7,9	1,3	-1,9
Mortality rate	0,0018	0,0017	0,0010	0,0008	0,0006
Observations	5,484,185	5,476,532	5,471,145	5,477,554	5,471,973
R-squared	0.00	0.00	0.00	0.00	0.00

Notes: This table presents the effect of regional unemployment rates on the probability of belonging to a short-term income quintile for the period 1983 to 2000, estimated with a linear probability model. Short-term income is defined as the one year lagged disposable family income. The models include controls for the year and regional fixed-effects, and regional time-trends. Robust standard errors are shown in parentheses.*** p<0.01, ** p<0.05 and * p<0.1 denote significance level.

Table A3. The Impact of the Business Cycle on Income-Related Inequality in Mortality for the Period 1979-2000.

	Short-term			Long-term		
	ECI	ARCI	ARCI	ECI	ARCI	SRCI
Regional unemployment rates	-0,00194	0,107	-0,00048	-0,00319	0,215	-0,0008
	(0,00307)	(0,599)	(0,00077)	(0,00361)	(0,645)	(0,0009)
Constant	-0,00095***	-0,218***	-0,00024***	-0,00162***	-0,371***	-0,0004***
	(0,00014)	(0,0274)	(0,00004)	(0,00016)	(0,0295)	(0,00004)
Observations	33,292,486	33,292,486	33,292,486	33,292,486	33,292,486	33,292,486

Notes: This table presents the effect of regional unemployment rates on male income-related inequality in mortality using the Recentered Influence Function (RIF) of the Erreygers Concentration Index (ECI), the Attainment-Relative (ARCI) and the Shortfall-Relative (SRCI) Concentration Index, as the dependent variable, respectively, in an OLS-regression for the period 1979 to 2000. The RIFs are calculated yearly at the national level using the individual mean of the age and year standardized lagged disposable family income as the ranking variable in the long-term analysis. The models include controls for the year and regional fixed-effects, and regional time-trends. *** p<0.01, ** p<0.05 and * p<0.1 denote significance level.