# Inequality in current and lifetime income

Rolf Aaberge, Magne Mogstad

This is a post-peer-review version of an article published in the journal *Social Choice and Welfare,* made available in accordance with the copyright policy of the journal publisher *Springer*. It may contain minor differences from the journal's pdf-version.

The final authenticated version is available at:

Aaberge, R., & Mogstad, M. (2015). Inequality in current and lifetime income. *Social Choice and Welfare*, 44(2), 217-230. https://doi.org/10.1007/s00355-014-0838-3

Statistisk sentralbyrå
Statistics Norway

# Inequality in Current and Lifetime Income

Rolf Aaberge
Research Department, Statistics Norway,
roa@ssb.no

Magne Mogstad
University of Chicago & Statistics Norway,
magne.mogstad@gmail.com

July 2014

**Abstract:** To gauge inequality in living standards, the distribution of lifetime income is likely to be more relevant than the distribution of current income. Yet, empirical studies of income inequality are typically based on observations of income for one or a few years. In this paper, we exploit a unique data set with nearly career-long income histories to assess the role of so-called life-cycle bias in empirical analysis of income inequality that uses current income variables as proxies for lifetime income. We find evidence of substantial life-cycle bias in estimates of inequality based on current income. One implication is that cross-sectional estimates of income inequality are likely to be sensitive to the age composition of the sample. A decomposition of the life-cycle bias into income mobility and heterogeneous profiles reveal the importance of two explanations that have been put forth to explain the disagreement between current and lifetime inequality.

**Keywords:** Income inequality, lifetime income, current income, life-cycle

**JEL codes:** D31; I31

**Acknowledgements:** We thank two anonymous referees, the editor and Anders Björklund for helpful comments.

## 1. Introduction

More than half a century ago, Friedman (1957, p38) stated that "*the identification of low measured income with 'poor' and high measured income with 'rich' is justified only if measured income can be regarded as an estimate of expected income over a lifetime or a large fraction thereof"*. Because individuals can borrow and smooth out life-cycle changes in income, he argued that the standard of living in any year depends more on lifetime income than on that year's income. A large body of empirical evidence support this conjecture (see e.g. Blundell, R., L. Pistaferri, and I. Preston, 2008). To gauge inequality in living standards, the distribution of lifetime income is therefore likely to be more relevant than the distribution of current income.

Yet, contemporary empirical studies of income inequality are typically based on observations of income for one or a few years. This empirical simplification is due to the simple fact that researchers seldom have access to data on long-run or lifetime income.[1] To provide evidence on inequality in current and lifetime income, we use a unique data set covering the entire Norwegian population with nearly career-long income histories for certain cohorts. We use this data set to assess the role of so-called life-cycle bias in empirical analysis of income inequality that uses current income variables as proxies for lifetime income.

The insights from our analysis can be summarized by two broad conclusions. First, we find evidence of substantial life-cycle bias in estimates of inequality based on current income. The life-cycle bias is minimized when individuals' incomes are measured at age 34-35. Later in the life, the bias increases steadily and inequality in current income at age 60 is more than twice the inequality in lifetime income. One implication is that cross-sectional estimates of income inequality are likely to be sensitive to the age composition of the sample. As a consequence, it is necessary to pay close attention to differences in age composition when comparing cross-sectional estimates of income inequality across countries, subgroups, or time.

Second, inequality in lifetime income is much lower that what cross-sectional estimates of inequality suggest. This means that we may need to reconsider how unequal individuals' living standard actually is. To understand better why inequality in current income exceeds inequality in lifetime income, we decompose the life-cycle bias into two components:

[1] There are, however, a few studies that use income data over a longer period. A notable example is Björklund (1993), who uses a long, panel data on income from Sweden to compare the distributions of annual and long-term income. He finds that the inequality in long-run income is around 35 to 40 percent lower than in the cross-sections of annual income. Burkhauser and Poupore (1997) report reductions in income inequality indices of about 15 to 20 percent when the accounting period was extended from 1 to 6 years. Using income data for a whole decade, both Gittleman and Joyce (1999) and Aaberge et al. (2002) find that cross-sectional inequality indices overstate inequality in long-run income by 25 to 30 percent. The focus of

income mobility (i.e. changes over time in individuals' positions in the current income distributions) and heterogeneous (but non-intersecting) age–income profiles. We find that income mobility reduces inequality in lifetime income by about 25 percent, while heterogeneous age–income profiles contributes to upward (downward) life-cycle bias in estimates of income inequality when current incomes are measured late (early) in the working lifespan. Overall, the equalizing effect of income mobility tends to dominate, implying that inequality in current income exceeds inequality in lifetime income even when current incomes are measured early in individuals' careers.

Our paper complements a growing literature in economics aimed at providing better-informed analyses of estimation biases in a wide range of research that uses current income variables as proxies for long-run income. Closely related to our paper is Haider and Solon (2006), who demonstrate that the association between current and lifetime income varies systematically over the life-cycle. As a consequence, regression models of intergenerational income mobility that use current income as a proxy for lifetime income will produce inconsistent estimates of the intergenerational elasticity. The inconsistency varies as a function of the age at which current income is observed, and is therefore called life-cycle bias.[2]

Our point of departure is that life-cycle variation in the association between current and lifetime income, might not be representative for life-cycle variation in the association between *inequality* in current and lifetime income: The former depends on the covariance between individuals' current and lifetime income; the latter depends on how the covariance between individuals' incomes and their position in the distribution of income varies with the age at which current income is measured. Our empirical analysis demonstrates that the essential features of life-cycle bias in income inequality cannot be revealed from the life-cycle association between the levels of lifetime and current income.

Our paper also relates to a literature on the modeling of labor income dynamics (see e.g. Meghir and Pistaferri, 2011). The typical specifications in this literature drops individuals not working over a whole year and works with residuals from regressions of log labor income on calendar time dummies and individual characteristics (such as education and age). Their analyses, therefore, examines how the within group inequality in log labor income of workers varies over time and across the life-cycle. The aim is to decompose the variance of residual log labor income into components like transitory and permanent shocks and heterogeneous

these studies is how income mobility reduces inequality in long-run income, and they do not study the issue of life-cycle bias in studies of inequality that use current income as a proxy for lifetime income.

[2] The empirical analysis of Haider and Solon has been replicated and extended for Sweden (Böhlmark and Lindquist, 2006),

profiles.[3] By comparison, we are interested in overall inequality in current and lifetime income, and therefore, do not eliminate between-group income differences. Nor do we restrict the sample to individuals who work.

This paper unfolds as follows. Section 2 presents a conceptual framework that relates inequality in current and lifetime income and illustrates the possible role of life-cycle bias. Section 3 describes our data and reports summary statistics. Section 4 provides evidence on inequality in lifetime and current income, before assessing the life-cycle bias. Section 5 concludes.

## 2. Conceptual framework

This section uses a framework of compensating differences, originally proposed by Mincer (1958), to relate inequality in distribution of current and lifetime income and illustrate the possible role of life-cycle bias.

Following Willis and Rosen (1979), suppose that individuals choose between two levels of schooling, labeled college (A) and high school (B), to maximize the present value of lifetime income. Assume that credit markets are perfect and the environment is perfectly certain, but occupations differ in the amount of schooling required. If an individual chooses college, his or her current income stream is

(2.1) $$y_{A,t} = \begin{cases} 0, & t \le s \\ \bar{y}_A e^{\gamma_A (t-s)}, & t > s \end{cases} ,$$

where $s$ is the number of years it takes to get a college degree, $t$ represents age (measured as years since high-school graduation), $\bar{y}_A$ is initial income, and $\gamma_A$ is the growth rate in income. If the individual chooses high school, his or her current income stream is

(2.2) $$y_{B,t} = \bar{y}_B e^{\gamma_B t}, t \ge 0,$$

where $\bar{y}_B$ is the initial income and $\gamma_A$ is the growth rate in income.

Germany (Brenner, 2010), and Norway (Nilsen et al., 2012). See also the discussion in Nybom and Stuhler (2014).

[3] Because of the small scale of the U.S. panel surveys, much of the early research relied on simple models that impose economically implausible restrictions (see the discussions in Baker and Solon, 2003). Using rich administrative data, Haider (2001), Baker and Solon (2003), Moffitt and Gottschalk (2012), and Blundell, Graber and Mogstad (2014) go beyond earlier models by allowing for key aspects in the evolution of labour income over time and across the life-cycle.

Suppose that additional schooling entails opportunity costs in the form of foregone income (but no direct cost such as tuition). Assume an infinite horizon, and an exogenously determined interest rate *r*, with $r > \gamma_A > \gamma_B > 0$. Then, the present value of income is

(2.3) $$Y_A = \int_0^\infty y_{A,t} e^{-rt} dt = \frac{\bar{y}_A e^{-rs}}{r - \gamma_A}$$

if college is chosen, and

(2.4) $$Y_B = \int_0^\infty y_{B,t} e^{-rt} dt = \frac{\bar{y}_B}{r - \gamma_B}$$

if high school is chosen. To induce a worker to choose college, foregone income while in school must be compensated by higher future income, such that $Y_A > Y_B$. In the long-run competitive equilibrium, the relationship between lifetime income and schooling is such that: (i) the supply and demand for workers of each schooling level are equated, and (ii) no worker wishes to alter his or her schooling level.

In the basic framework of compensating differences, individuals are ex ante identical. Suppose that inequality is measured by the Gini coefficient (*G*), and let $p \in (0,1)$ denote the equilibrium proportion of workers with high school. Then, income inequality in equilibrium is given by

(2.5) $$G_t = \frac{p(1-p)\left|y_{A,t} - y_{B,t}\right|}{p y_{B,t} + (1-p) y_{A,t}}$$

if current income is measured at age *t*, and by

(2.6) $$G = \frac{p(1-p)\left|Y_A - Y_B\right|}{p Y_B + (1-p) Y_A}$$

if lifetime income forms the basis for the analysis.

In the spirit of Haider and Solon (2006), we define $G_t - G$ as the *life-cycle bias* in income inequality from using current income at age *t* as a proxy for lifetime income. Equilibrium equates lifetime incomes such that individuals are indifferent between schooling levels, implying that $G$ is equal to zero. Since foregone income while in school must be compensated, the equilibrium $G_t$ will generally be non-zero and vary as a function of the age at which $y_t$ is observed. In our case, the age at which $G_t$ is zero and life-cycle bias is eliminated is uniquely given by

(2.7) $$t^* = \frac{\log \bar{y}_B - \log \bar{y}_A + s\gamma_A}{\gamma_A - \gamma_B},$$

which is obtained by inserting (2.1) and (2.2) in $y_{A,t} = y_{B,t}$.

More realistic models of income allow for ex ante heterogeneous individuals, such as in initial wages, growth rates, and the interest rate (see e.g. Willis and Rosen, 1979). Even so, the crucial insight of the basic framework of compensating differences still applies: to induce a worker to undertake additional schooling, foregone income while in school must be compensated by higher future income. This should generate changes in the population variance of income around the central tendency of income growth, causing life-cycle bias in analysis of inequality using cross-section data. Indeed, empirical evidence suggest the income profiles of education groups fan out with age, so that the earnings premiums to education increases over the life-cycle (see e.g. Bhuller et al., 2014).

To circumvent the issue of life-cycle bias, the data used in analysis of income inequality would ideally consist of complete longitudinal life histories of income. Unfortunately, such ideal data are seldom available. In the remainder of the paper, we provide evidence on inequality in current and lifetime income from nearly career-long income histories, and assess the role of life-cycle bias in empirical analysis of income inequality. Our analysis also provides an estimate of $t^*$. Restricting the sample to individuals around this age may be used as an imperfect, short-cut method for approximating inequality in lifetime from cross-section data.

## 3. Data and definitions

### 3.1 Data

Our empirical analysis uses a longitudinal dataset containing records for every Norwegian from 1967 to 2006. The variables captured in this dataset include demographic information (sex, year of birth) and income. We focus on the 1942-1944 cohorts in order to ensure availability of data on income for more or less the entire working lifespan. In particular, these cohorts are between 23 and 25 years old in 1967 and between 62 and 64 in 2006.[4] Our analytical sample is restricted to males, given their role of breadwinner and primary wage-earner for these cohorts. Also, we exclude individuals whose information on annual income is missing. The final sample used in the analysis consists of 51 552 individuals.

Our measure of income includes all taxable income after deductions. Taxable income is the sum of pretax market income (from wages and self-employment), cash benefits (such as unemployment benefits, disability benefits, and family benefits), and capital income. We use the consumer price index to make incomes from different years comparable.

Following Haider and Solon (2006), our measure of *lifetime income* is the annuity value of the discounted sum of real income

$$Y = \frac{y_T + \sum_{t=1}^{T-1} y_t \prod_{j=1+t}^{T} (1+r_j)}{1 + \sum_{t=1}^{T-1} \prod_{j=1+t}^{T} (1+r_j)}, \tag{3.1}$$

where $r_t$ denotes the real interest rates on income-transfers from period *t-1* to *t*.[5] We define *current income* as the average real income over a two-year period. The use of two years of income observations is a common procedure to minimize the contamination of inequality estimates by transitory income variation and measurement error (see e.g. Jenkins and Van Kerm, 2006). In our setting, "true" measurement error is unlikely to be important. In particular, there is little scope for reporting or recollection errors; the data come from individual tax records with detailed information about the different sources of income. Indeed, the coverage and reliability of Norwegian registry data on income are considered to be exceptional (Atkinson, Rainwater, and Smeeding, 1995).

It should be noted that the Norwegian income data also have other advantages over those available in many other countries. First, there is no attrition from the original sample because it is not necessary to ask permission from individuals to access their tax records. In Norway, these records are in the public domain. Second, our income data pertain to all individuals, and not only to jobs covered by social security. Third, we have nearly career-long income histories for certain cohorts, and do not need to extrapolate the income profiles to ages not observed in the data. And fourth, our income measure is neither top-coded nor bottom-coded.

[4] Although the formal retirement age is 67 years, many individuals retire around age 65.

### 3.2 Descriptive statistics

Table 1 reports descriptive statistics. The first two columns display the mean and standard deviation in current and lifetime income for the whole sample. We can see that average income increases over the life cycle. Average current income is most similar to average lifetime income when individuals are in their 40s. The increase in average current income over the life cycle is accompanied by an increase in the variance of current income. This is an important observation, because life-cycle variation in inequality is due to changes in income variation around the central tendency of income growth.

In the last three columns of Table 1, we partition the sample into three groups according to the distribution of lifetime income. We can see that individuals with high lifetime income (cf. Column 5) experience, on average, rapid income growth over the life-cycle as compared to individuals with medium (cf. Column 4) or low (cf. Column 3) lifetime income. The life-cycle bias in inequality based on current income at a particular age depends on how well the differences in current income approximate the differences in lifetime income. The table suggests that the income gap late in individuals' careers tends to overstate the lifetime income gap. Taken at face value and assuming no income mobility, this would indicate that there is an upward (downward) life-cycle bias in estimates of income inequality when incomes are measured late (early) in the working lifespan.

**Table 1. Descriptive statistics for current and lifetime income (NOK)**

| | Full sample | | ——— Subsample by lifetime income ——— | | |
|---|---|---|---|---|---|
| | Mean | St. Dev. | 1st quartile group | 2nd and 3rd quartile group | 4th quartile group |
| Current income | | | | | |
| Age 26-27 | 177 883 | 64 753 | 138 304 | 177 659 | 217 912 |
| Age 38-39 | 245 814 | 101 431 | 167 555 | 236 789 | 342 124 |
| Age 50-51 | 272 246 | 423 089 | 145 380 | 233 691 | 476 223 |
| Age 62-63 | 330 922 | 689 784 | 167 061 | 261 474 | 633 681 |
| Lifetime income | 249 523 | 64 753 | 160 577 | 232 852 | 371 811 |
| Observations | 51 552 | | 12 888 | 25 776 | 12 888 |

*Notes*: The sample consists of males born 1948-1950. Current income is defined as the average income over the two-year period. Lifetime income is defined as the annuitized value of real income from age 24 to 63. Age refers to the 1949 cohort.

[5] The annual real interest rates on borrowing and savings are computed from Norwegian official statistics on interest rates on loans and deposits in commercial banks over the period 1967–2006.

### 3.3 Inequality measures

To achieve rankings of intersecting Lorenz curves, many empirical studies employ the Gini coefficient. To examine the extent to which the empirical results depend on the choice of inequality measure, the conventional approach is to complement the Gini coefficient with measures from the Atkinson or the Theil family. However, the Gini coefficient and measures from the Atkinson or the Theil family have distinct theoretical foundations which make it inherently difficult to evaluate their capacities as complimentary measures of inequality.

As demonstrated by Aaberge (2000, 2007), an alternative approach for examining inequality in the distribution of income is to rely on the moments of the scaled conditional mean curve, defined by

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

where $F$ is the observed cumulative distribution function with mean $\mu$, and $F^{-1}$ its left inverse.[6] Specifically, the first, second, and third order moments of the scaled conditional mean curve prove to make up a fairly good summarization of the conditional mean curve as well as the Lorenz curve. The $k^{th}$ order moment of the scaled conditional mean curve for $F$, denoted $C_k(F)$, is defined by

$$C_k(F) = k\int_0^1 u^{k-1}(1 - N(u))du, \quad k = 1,2,\ldots. \tag{3.2}$$

Aaberge (2007) characterizes the complementary nature of the first three moments of the conditional mean curve; they all satisfy the Pigou-Dalton transfer principle but differ in their sensitivity to inequality in the lower, central, and upper part of the distribution. The first moment ($C_1$) is equal to the Bonferroni coefficient, which is particularly sensitive to inequality in the lower part of the distribution. The second moment is equal to the Gini coefficient ($C_2 = G$), which is known to assign a lot of weights to inequality in the central part of the distribution. The third moment is an inequality measure ($C_3$) that is particularly sensitive to inequality in the upper part of the income distribution.

[6] For a given $u$, $N(u)$ is the ratio of the mean income of the poorest 100u per cent of the population and the overall mean. By inserting for the Lorenz curve in (3.2) it follows straightforwardly that the scaled conditional mean curve is a representation of inequality that is equivalent to the Lorenz curve.

We will examine the sensitivity of the empirical results to the choice of inequality measure by complementing the information provided by G with its two close relatives $C_1$ and $C_3$.[7] Hence, we meet the most common criticism of the Gini-coefficient, namely that it is insensitive to inequality in the tails of the distribution (see e.g. Wiles, 1974).

### 3.4 Mobility measures

To understand better why inequality in current incomes differs from inequality in lifetime incomes, we need an appropriate measure of income mobility. Let $F_Y$ and $F_Z$ denote the distribution of the observed lifetime income $Y$ and the distribution of the reference lifetime income Z in the hypothetical situation of no income mobility. The latter distribution is formed by assigning the lowest income in every period to the poorest individual in the first period, the second lowest to the second poorest, and so on. Accordingly, the design of the distribution of Z does not alter the marginal period-specific distributions. Since $F_Y$ can be attained from $F_Z$ by a sequence of period-specific Pigou-Dalton income transfers, we know that

$$C_k(F_Z) \geq C_k(F_Y)$$

with strict equality insofar $F_Y$ is identical to $F_Z$.

In line with Shorrocks (1978), we can then measure income mobility as

$$M_k(F) = \frac{C_k(F_Z) - C_k(F_Y)}{C_k(F_Z)} = \frac{\sum_{t=1}^{T} b_t \frac{\mu_t}{\mu} C_k(F_{Y_t}) - C_k(F_Y)}{\sum_{t=1}^{T} b_t \frac{\mu_t}{\mu} C_k(F_{Y_t})} \tag{3.3}$$

where

$$b_t = \frac{\prod_{j=t+1}^{T}\left(1+r_j\right)}{1+\sum_{s=1}^{T-1}\prod_{j=s+1}^{T}\left(1+r_j\right)}, \quad t = 1,2,\ldots,T-1;\, b_T = 1.$$

It follows straightforward that $0 \leq M_k \leq 1$. Mobility measures based on (3.3) can be interpreted as the relative reduction in inequality in lifetime income due to changes in

[7] Since $C_1$, G, and $C_3$ have a common theoretical foundation and complement each other with regard to sensitivity to changes in the lower and upper part of the income distribution, Aaberge (2007) calls them Gini's Nuclear Family.

individuals' positions in the short-term distributions of income. The state of no mobility ($M_k = 0$) occurs when individuals' positions in the short-term income distributions are constant over time. By comparison, the mobility measures takes the maximum value of one ($M_k$ = 1) when changes in individuals' positions in the short-term income distributions generate complete equality in the distribution of lifetime incomes.

## 4. Empirical results

This section begins by reporting estimates of inequality in lifetime and current income. This allows us to assess the life-cycle profile in income inequality, and to identify the age at which life-cycle bias is minimized. We next decompose the life-cycle bias into income mobility and heterogeneous income profiles. We conclude this section by demonstrating how the essential features of life-cycle bias in inequality cannot be revealed from the life-cycle association between the levels of lifetime and current income.

### 4.1. Inequality in current and lifetime income

Figure 1 graphs the estimated inequality in lifetime income and current income, age 24–63. There are three clear patterns in our results, independent on whether the inequality measure focuses on income differences that take place in the lower ($C_1$), middle ($C_2 = G)$), or upper ($C_3$) part of the income distribution.

First, we see that inequality in current income is quite stable until individuals are in their late 30s, after which inequality increases steadily. This pattern reflects that individuals with high lifetime income experience more income growth over the life-cycle.

Second, we find evidence of systematic life-cycle bias in estimates of inequality based on current income. Recall that $C_k(F_{Y_t}) - C_k(F_Y)$ measures the *life-cycle bias* in income inequality from using current income at age *t* as a proxy for lifetime income. Figure 1 shows that the life-cycle bias is minimized when individuals' incomes are measured at age 34-35. Later in the life, the bias increases steadily and inequality in current income at age 60 is more than twice the inequality in lifetime income.

Third, inequality in lifetime income is much lower than what cross-sectional estimates of inequality suggest. For example, Aaberge and Mogstad (2011) study the evolution of income inequality in Norway over the period 1967-2007. In each year, their sample consists of the entire population of males aged 20–65. The cross-sectional estimates of the Gini coefficient increase from around .26 in 1967 to about .36 in 2006. In contrast, Figure 1 shows that the inequality in lifetime income over this period is less than .19.

### 4.2. Decomposition of life-cycle bias

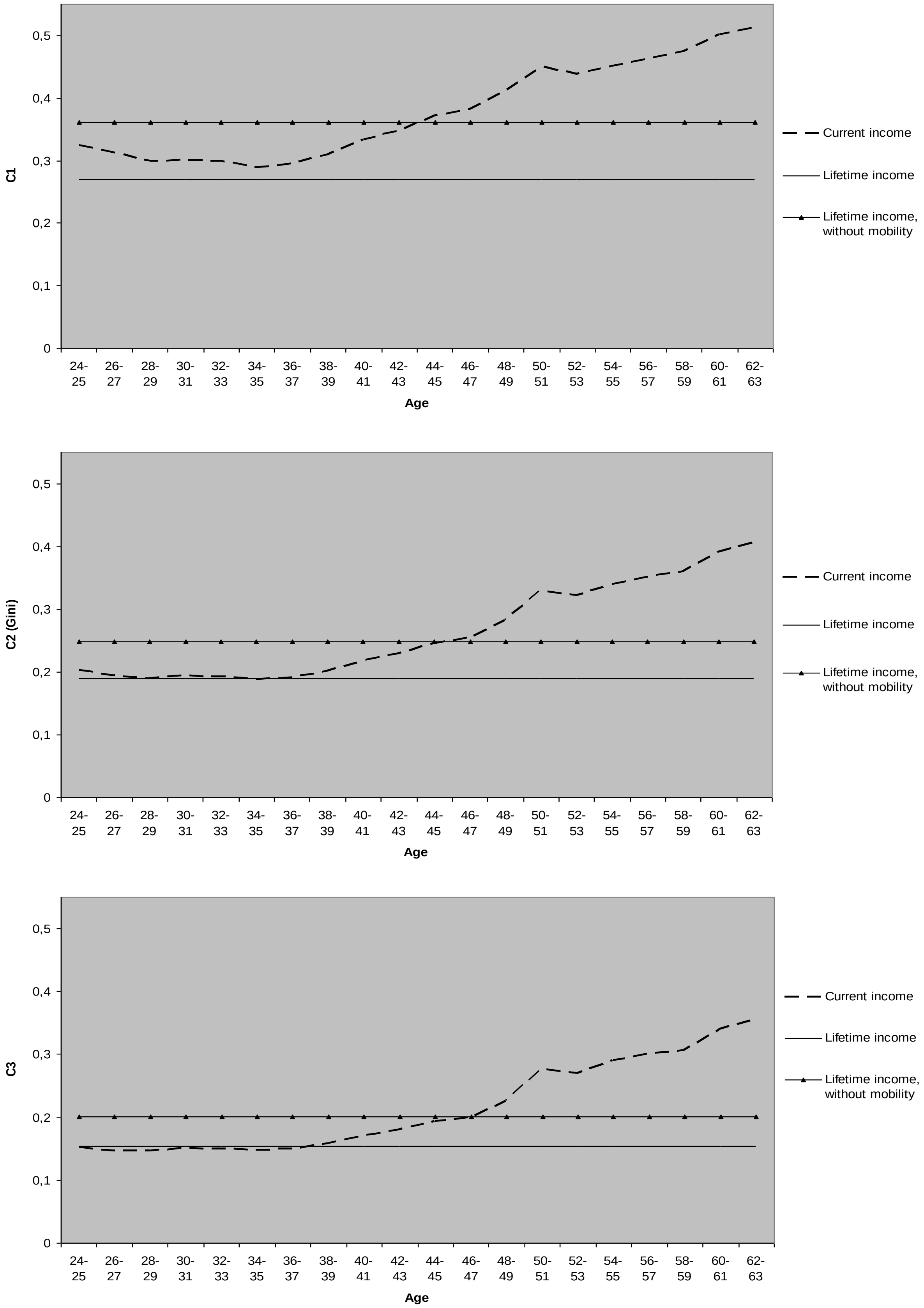
To understand why inequality in current incomes differs from inequality in lifetime incomes, we decompose the life-cycle bias into two components: income mobility (i.e. changes over time in individuals' positions in the current income distributions) and heterogeneous (but non-intersecting) age–income profiles. From (3.3), we get that

$$C_k(F_{Y_t}) - C_k(F_Y) = C_k(F_{Y_t}) - C_k(F_Z)\left(1 - M\right). \tag{4.1}$$

Recall that $C_k(F_Z)$ can be expressed as a weighted average of inequality in current income over the life-cycle. In the absence of income mobility (M=0), $C_k(F_{Y_t}) - C_k(F_Y)$ will therefore be either zero at all ages or vary in sign over the life-cycle. Because of income mobility, however, inequality in current income may exceed inequality in lifetime income at all ages.

Figure 1 displays $C_k(F_Y)$, $C_k(F_Z)$, and $C_k(F_{Y_t})$ *for* $k = 1, 2, 3$. The figure shows that income mobility reduces inequality in lifetime income by about 25 percent, while heterogeneous (but non-intersecting) age–income profiles contributes to upward (downward) life-cycle bias in estimates of income inequality when current incomes are measured late (early) in the working lifespan. Overall, the equalizing effect of income mobility tends to dominate, implying that inequality in current income usually exceeds inequality in lifetime income even when current incomes are measured early in individuals' careers.

**Figure 1. Inequality estimates in current and lifetime income**




*Note:* The sample consists of males born 1948-1950. Current income is defined as the average income over the two-year period. Lifetime income is defined as the annuitized value of real income from age 24 to 63. Age refers to the 1949 cohort. Lifetime income without mobility is defined as assigning the lowest income in every period to the poorest individual in the first period, the second lowest to the second poorest, and so on.

### 4.3 Life-cycle variation in the association between current and lifetime income

To evaluate the appropriateness of the textbook errors-in-variables model, Haider and Solon (2006) analyzed life-cycle variation in the association between current and lifetime income. They found that the slope coefficient in a regression of current income on lifetime income

$$\lambda_t = \frac{cov(y_t, Y)}{var(Y)},$$

is generally not equal to one, as the textbook errors-in-variables model assumes, varying instead systematically over the life cycle. An important implication is that using current income as a proxy for lifetime income can generate substantial life-cycle bias in regression analysis. Haider and Solon propose a possible remedy: Under the assumptions of the generalized errors-in-variables model, life-cycle bias is eliminated by using current income around the age when $\lambda_t$ equals one. The procedure of measuring income around a certain age has over the last few years been much used to correct for life-cycle bias applied research, especially in analysis of intergenerational income mobility.[8]

We begin by redoing the analysis of Haider and Solon in the Norwegian context. Our estimates of $\lambda_t$ start out at a value less than one at the outset of the career, then increase over the life cycle, and exceed one later in the career. In our data set, $\lambda_t$ is closest to one around when individuals' incomes are measured in their late 40s.[9] However, this finding does not justify using current income in the late 40s as a proxy for lifetime income in the measurement of inequality. To see this, note that the life-cycle variation in the association between the Gini coefficient in current and lifetime income can be expressed as

$$G_t - G = \frac{2cov[y_t, F(y_t)]}{E(y_t)} - \frac{2cov[Y, F(Y)]}{E(Y)}.$$

In general, $\lambda_t$ = 1 does not imply $G_t = G$. The former depends on the covariance between individuals' current and lifetime income; the latter depends on how the covariance between

[8] For studies of intergenerational income mobility taking this approach to correct for life-cycle bias, see the review of Black and Devereux (2011).

[9] In comparison, Haider and Solon (2006) find that $\lambda_t$ comes close to 1 when individuals are in their early 40s (and mid 30s), whereas Bohlmark and Lindquist (2006) report that $\lambda_t$ is approximately 1 when individuals are aged 46-53 (and around age 33). It should be noted, however, that these two studies measure income in logs rather than levels. When measuring income in logs, our estimate of $\lambda_t$ is equal to 1 around age 40.

individuals' incomes and their position in the distribution of income varies over the life-cycle. Figure 1 demonstrates that this difference is it not merely a theoretical peculiarity, but empirically important: Life-cycle bias in inequality is minimized at age 34-35, whereas current income at the age when $\lambda_t$ is close to one grossly overstates the extent of inequality in lifetime income

## 5. Conclusion

Research on the income inequality has a long history in economics. In particular, considerable effort has been directed towards examining how cross-sectional estimates of inequality may overstate inequality in long-run incomes because of income mobility. By way of comparison, much less attention has been devoted to the life-cycle bias that may arise from the widespread use of current income as a proxy for lifetime income in measurement of inequality.

This paper helps to close that gap by exploiting a unique data set with nearly career-long income histories to assess the role of life-cycle bias in empirical analysis of income inequality. We find evidence of substantial life-cycle bias in estimates of inequality based on current income. One implication is that cross-sectional estimates of income inequality are likely to be sensitive to the age composition of the sample. As a consequence, it is necessary to pay close attention to differences in age composition when comparing cross-sectional estimates of income inequality across countries, subgroups, or time. A decomposition of the life-cycle bias into income mobility and heterogeneous profiles reveal the importance of two explanations that have been put forth to explain the disagreement between current and lifetime inequality.

Our findings serve to highlight the importance of addressing life-cycle bias in empirical analysis of income inequality. At the same time, they raise a number of questions. To what extent does the relationship between inequality in current and lifetime income depend on the choice of income concept? For example, what happens if one subtracts taxes or adds in income sources of spouses? How does the pattern of life-cycle bias vary over time or across countries? What is the relative importance of calendar time effects and age effects in the life-cycle profiles in inequality? Answers to these questions are key to understand the breadth and nature of life-cycle bias in analysis of economic inequality.

## 6. References

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

Aaberge, R. (2007): “Gini’s Nuclear Family,” *The Journal of Economic Inequality,* **5,** 305-322.

Aaberge, R., A. Björklund, M. Palme, P. Pedersen, N. Smith, T. Wennemo (2002): “Income Inequality and Income Mobility in the Scandinavian Countries Compared to the United States,” *Review of Income and Wealth,* **48**, 443-469.

Aaberge, R. and M. Mogstad. (2011): “Robust Inequality Comparisons,” *The Journal of Economic Inequality,* **9,** 353-371.

Atkinson, A., L. Rainwater, and T. Smeeding (1995): “Income Distribution in OECD Countries: Evidence from the Luxembourg Income Study," Report, OECD.

Baker, M. and G. Solon (2003): “Earnings Dynamics and Inequality among Canadian Men, 1976-1992: Evidence from Longitudinal Income Tax Records”, Journal of Labor Economics, **21**, 267-288.

Blundell, R., L. Pistaferri, and I. Preston (2008): “Consumption inequality and Partial Insurance," American Economic Review, **98**, 1887-1921.

Blundell, R., M. Graber, and M. Mogstad (2014): “Labor Income Dynamics and the Insurance from Taxes , Transfers , and the Family“, forthcoming in Journal of Public Economics.

Bhuller, M., M. Mogstad, and K. Salvanes (2014): “Life Cycle Earnings, Education Premiums and Internal Rate of Returns”, Working Paper.

Björklund, A. (1993): “A Comparison between Actual Distributions of Annual and Lifetime Income: Sweden 1951-89,” *Review of Income and Wealth,* **39,** 377-386.

Black, S. and Devereux, P. (2011): “Recent Developments in Intergenerational Mobility.” In O. Ashenfelter and D. Card (eds): *Handbook of Labor Economics Vol 4*. Elsevier.

Brenner, J. (2010): “Life-Cycle Variations in the Association between Current and Lifetime Earnings: Evidence for German Natives and Guest Workers,” *Labour Economics,* **17,** 392-406.

Burkhauser, R. V. and Poupore, J. G. (1997): “A Cross-National Comparison of Permanent Inequality in the United States and Germany,” *The Review of Economics and Statistics,* **79,** 10-17.

Böhlmark, A. and Lindquist, M. J. (2006): “Life-Cycle Variations in the Association between Current and Lifetime Income: Replication and Extension for Sweden,” *Journal of Labor Economics,* **24,** 879-900.

Friedman, M. (1957): “A Theory of the Consumption Function,” Princeton University Press, Princeton.

Gittleman, M. and Joyce, M. (1999): “Have Family Income Mobility Patterns Changed?” *Demography,* **36,** 299-314.

Haider, S. (2001): “Earnings Instability and Earnings Inequality of Males in the United States: 1967-1991”, Journal of Labor Economics, **19**, 799—836.

Haider, S. and Solon, G. (2006): “Life-Cycle Variation in the Association between Current and Lifetime Earnings,” *American Economic Review,* **96**, 1308-1320.

Jenkins, S. P. and Van Kerm, P. (2006): “Trends in Income Inequality, Pro-Poor Income Growth and Income Mobility,” *Oxford Economic Papers,* **58,** 531-548.

Nybom, M, and J. Stuhler (2014): “Heterogeneous Income Profiles and Life-Cycle Bias in Intergenerational Mobility Estimation”, Working Paper.

Meghir, C. and L. Pistaferri (2011): “Earnings, Consumption and Lifecycle Choices”. In O. Ashenfelter and D. Card (eds): *Handbook of Labor Economics Vol 4*. Elsevier.

Mincer, J. (1958): “Investment in Human Capital and Personal Income Distribution,” *The Journal of Political Economy,* **66,** 281-302.

Moffitt, R. and P. Gottschalk (2012): “Trends in the Transitory Variance of Male Earnings: Methods and Evidence“, Journal of Human Resources, **47**, 204-236

Nilsen, Ø, K. Vaage, A. Aakvik, and K. Jacobsen (2012): “Intergenerational Earnings Mobility Revisited: Estimates Based on Lifetime Earnings,” *Scandinavian Journal of Economics,* **114,** 1-23.

Shorrocks, A. (1978), “Income Inequality and Income Mobility”, *Journal of Economic Theory*, 19*,* 376-93.

Wiles, P. (1974): “Distribution of Income, East and West,” North Holland Publishing Company, Amsterdam.

Willis, R. J. and Rosen, S. (1979): “Education and Self-Selection,” *Journal of Political Economy,* **87,** 7-36.