Human Capital Inequality: Empirical Evidence†

Brant Abbott

Giovanni Gallipoli

†We acknowledge support through an Insight grant from the SSHRC of Canada. We are grateful to Sarah O’Brien, Robin Li, Pietro Montanarella and Lily Suh for excellent research assistance.
Dispersion of Economic Outcomes across Households

Persistent differences in economic outcomes, such as income, wealth, employment and consumption, have received continuing attention in the academic debate. Part of this attention is motivated by the well documented increase in cross-sectional economic inequality that began in the 1980s. Alternative measures of inequality, such as variances, inter-percentile ranges and concentration at the top, all suggest that this increase has occurred across a range of measurable outcomes (earnings, wealth, health).

Wealth inequality, in particular, has received considerable attention, with mounting evidence of steady and economically meaningful changes in the concentration of wealth ownership. By definition, wealth inequality captures disparity in the ownership of productive capital and other non-labor factors of production. In contrast, this article focuses on the distribution of human capital and its implications for the accrual of economic resources to individuals and households. Human capital inequality can be thought of as measuring disparity in the ownership of labor factors of production, which are usually compensated in the form of wage income. While not directly observable, the notion of a stock of human capital and its main properties have been formally spelled out in the original work of Gary Becker and coauthors (e.g. Becker, 1962, 1964).

Earnings inequality is tightly related to human capital inequality. However, earnings dispersion can only provide a partial and incomplete perspective on the underlying distribution of productive skills and on the income generated by way of them. This follows from the fact that, at any point in time, earnings dispersion captures an isolated snapshot of the prevailing disparity in payments to labour. In fact, one would rather learn about the distribution of the underlying stocks of human capital that generate income, and about their market value.

This tension between measures of earnings in a specific year and measures of long-term ability to earn has become a distinguishing feature of the literature on inequality and human capital. Despite its shortcomings, a fairly common way to gauge the distributional implications of human capital inequality is to examine the distribution of labor income. While it is not always obvious what accounts for returns to human capital, an established approach in the empirical literature is to decompose measured earnings into permanent and transitory components as in, for example, Gottschalk et al. (1994), Heathcote et al. (2010a) and recently Lochner and Shin (2014). The decomposition approach finds its theoretical motivation in the observation by Friedman (1957) that the underlying ability to earn associated with human capital must be related to the expected flow of income over an individual’s life. Because the permanent component of earnings observed in a given period should, by definition, be detectable in every
realization of an individual’s earnings process, one can regard an increase in the estimated variance of permanent earnings shocks as a proxy for the underlying changes in the value of human capital.

A second approach to the measurement of human capital inequality focuses on the lifetime present value of earnings. An early example of this is Lillard (1977), while more recent contributions are by Bowlus and Robin (2004) and Guvenen et al. (2017). Lifetime earnings are, by definition, an ex-post measure that is only observable at the end of an individual’s working lifetime. This measure captures how an individual’s position in the labour market changes over their lifetime through the evolution of their wage and employment status, rather than their position at a single point in time. One limitation of this approach is that it assigns a value based on one of the many possible realizations of human capital returns. That is, because it is impossible to re-run the working life of the same individual multiple times, this measure can only capture the value associated with the particular earnings realization that is observed for each individual. Arguably, this provides a partial view of human capital because it ignores the option value associated with alternative, but unobserved, potential earning paths that may be valuable ex-ante. Hence, ex-post lifetime earnings reflect both the genuine value of human capital and the impact of the particular realization of unpredictable shocks (luck): crucially these two components are not separately observable.

A different but related measure focuses on the ex-ante value of expected lifetime earnings, which differs from ex-post (realized) lifetime earnings insofar as they account for the value of yet-to-be realized pay-offs along different potential earning paths. Ex-ante expectations reflect how much an individual reasonably anticipates earning over the rest of their life based on their current stock of human capital, averaging over possible realizations of luck and other income shifters that may arise. The discounted value of different potential paths of future earnings can be computed using a risk-less discount factor as, for example, Jorgenson and Fraumeni (1989) and Cunha et al. (2005) do. This requires a way to estimate the distribution of future earnings outcomes. However, when valuing such outcomes it may be more appropriate to discount their value in a state-dependant fashion, as suggested in Huggett and Kaplan (2016) and Abbott and Gallipoli (2019). These more theory-focused approaches take estimates of the expectation of the present value of lifetime earnings as their yardstick measure of human capital, providing an indirect way to account for some of the key empirical issues highlighted above.

The empirical literature on economic inequality is so expansive that one cannot hope to summarize it all within a few pages. Instead, this paper focuses on a narrow aspect of inequality – namely human capital disparities – and attempt to provide an in-depth review of this particular dimension of inequality within a short article. The approach of this article is to dissect and
categorize existing evidence into different metrics of human capital, and then summarize and convey the critical information from each of those threads of the literature. While it is impossible to cover every article that might fit under each alternative measurement approach, this paper attempts to provide an overview of relevant and recent advances within those threads. Combining the evidence reported here with evidence on other dimensions of inequality, such as wealth and consumption, should allow the reader to obtain a more complete understanding of overall economic inequality.

Finally, in addition to reporting evidence on human capital inequality in the United States, this article also briefly outlines some international comparisons and provides a cursory discussion of ideas from the ever expanding literature on the determinants of human capital heterogeneity.

**Concepts of Human Capital and their Measurement**

While human capital is not directly observable, one can measure the distribution of returns to human capital utilization in the form of earnings. This allows researchers to draw inferences about the underlying distribution and value of human wealth. Moreover, given the multifaceted nature of human capital, one is also able to gauge variation in some of its most relevant features, such as educational achievement, health status and longevity. This section discusses relevant measurement issues and overview key results on the distribution of human capital.

**Earnings Inequality**

Earnings are observed at the individual level, but are sometimes analyzed at the household level. Earnings capture the part of income that is payment for labour. Thus earnings can be interpreted as the annual dividend paid by one’s human capital. Arguably, this dividend is not the best measure of human capital, but it is by far the most popular.

Some of the highest quality earnings data are sourced from government tax or social security data; however, such data cannot be freely accessed for analysis and are usually maintained in access-controlled research data centres. Earnings are also observable in survey data, such as the Census, though the reliability of such sources may be less than ideal.

One way to measure earnings inequality is to compute ratios of earnings of individuals at different percentiles of the income distribution. Researchers usually report the 90-10 ratio (90th percentile relative to 10th percentile), as well as the 90-50 and 50-10 ratio to further decompose what half of the distribution is driving changes in inequality. Such statistics are easily computed.
as one simply determines earnings at the relevant percentiles and takes their ratios.

Another common measure of earnings inequality is the top share of earnings. This is the share of total earnings in the economy accruing to those with the highest earnings, such as the highest 1% or 10% of earnings. Such statistics are also easy to compute as one determines the relevant percentile threshold (e.g. the 99th percentile) and then computes the total earnings of those above this threshold. Then, dividing this number by total earnings in the entire economy one arrives at the top-share statistic.

One of the most frequently cited measures of earnings inequality is the Gini coefficient. The interpretation of the Gini coefficient can be related to the Lorenz curve, which describes the proportion of economy-wide earnings attributed to those at, or below, a given percentile of the earnings distribution. When earnings are equally distributed (every person has identical earnings) the Gini coefficient is zero. In contrast, when earnings are perfectly unequal (one person gets everything, all others get nothing) the Gini coefficient is one. Earnings are a non-negative quantity, which ensures that the Gini coefficient is between zero and one. However other variables, such as net worth, can assume negative values and their distributions can therefore exhibit a Gini coefficient exceeding one.

A tremendous amount of data and measurement exists on income and earnings inequality in the United States. When analyzing inequality in human capital in the US one valuable data source is Social Security Administration (SSA) records, for which reported amounts are labor income and the unit of analysis is the individual. Another popular source for computing inequality facts is the IRS (U.S. Internal Revenue Service) database. However IRS records report total income and the unit of analysis is the tax-unit, which could be an individual or a couple.

An excellent description of facts and figures, based on SSA data, is in Kopczuk et al. (2010). These authors consider 80-50 and 50-20 ratios rather than the usual 90-50 and 50-10 ratios. They find that while the 80-50 ratio dipped after the Second World War (WWII) to as low as 1.5, a continuous steady rise in this ratio followed: the ratio climbed above 1.8 by the mid 2000s. This pattern reflects a steady increase in above-median inequality in the post-war era. The 50-20 ratio fell overall in this period, starting at about 2.5 at the end of WWII, and hovering around 2.2 by the mid-2000s, indicating a decrease in below-median inequality; however, this decline was not steady, with periods of increase occurring in the early 1980s and early 2000s. The overall inequality measure (the 80-20 ratio) showed a drop after WWII, and substantial increases thereafter, as the above-median increases in inequality outpaced the lower inequality in below-median earnings. These authors also report the time series of the Gini coefficient for earnings, which dropped to as low as about 0.34 in the early 1950s and then steadily rose to just under 0.45 by the mid-2000s. Figure 1 reproduces a figure from the original study of Kopczuk.
et al. (2010), which displays Gini coefficients for earnings by gender and over time. A similar
dynamic is observed in top earnings shares: the top 1% share of earnings dropped from 9.55% in
1939 to 5.92% in 1960, but then more than doubled to 12.28% by the mid 2000s. These
general patterns and magnitudes are consistent with many other studies and data sources, for
example Heathcote et al. (2010b) and Eckstein et al. (2004), who use Current Population Survey
data for similar time periods, or Rios-Rull and Kuhn (2013), who use data from the Survey of
Consumer Finances between 1989 and 2013.4.

Lastly, recent evidence has suggested that changing patterns of inequality are not a general
phenomenon, but rather are linked to growth of inequality within and across groups. For ex-
ample, Lemieux (2006c) shows that much of the previously discussed increases in inequality
are attributable to the increasing return to postsecondary education.5 In a related line of re-
search Song et al. (2015) find that most of the observed increase in earnings inequality can be
attributed to growing differences between firms in the wages they pay to workers. Similarly,
Barth et al. (2016) find that rising inequality can be partly attributed to increasing inequality
across industries.6

**Lifetime Earnings Inequality**

In theory, lifetime earnings correspond to the present discounted value of all labour-related
earnings over the course of one’s entire life. This concept is closer to the definition of human
capital than annual earnings is, as it captures the entire lifetime of dividends paid by one’s hu-
man capital. Importantly, these are the actual realized earnings of the individual, rather than
forecasted earnings, and so, to obtain a measure of this quantity, the individual’s entire lifetime
of earnings must be observed. In practice, lifetime earnings can be approximated by the present
discounted sum of annual earnings over some arbitrary period of time. As examples, Lillard
(1977) studies a narrow data set on post-schooling earnings, Guvenen et al. (2017) study exten-
sive administrative earning records for workers between ages 25 to 55. When the object of in-
terest is differences in lifetime income across cohorts, like in Guvenen et al. (2017), researchers
can avoid specifying a discount rate by assuming that all cohorts discount at the same rate. In
some instances, such as Bowlus and Robin (2004) and Low et al. (2010), theoretical models of
life-cycle earnings are used to interpret data panels and generate estimates of realized lifetime
earnings distributions. Once estimates of lifetime earnings are obtained for a cross-section of
individuals, researchers can compute the same measures of inequality described above in the
context of current (annual) earnings.
Guvenen et al. (2017) use 57 years of panel data (1957-2004) from the Social Security Administration to study the lifetime earnings of 27 birth cohorts. Each cohort is defined by the year that they turn 25, and the sample period spans 31 years. Annualized lifetime earnings are defined as the sum of real annual labour income between ages 25 and 55, divided by 31. Lifetime earnings of men and women are examined separately to reach two broad conclusions: first, men have experienced a stagnation in lifetime income while women have seen a rise in lifetime income; second, although lifetime income inequality within each gender group has risen, the closing of the gender gap has kept overall lifetime inequality nearly flat. Indeed, the 75-25 inter-percentile ratio in lifetime earnings was fairly steady (at about 3) for cohorts turning 25 between 1957 and 1983. However, looking separately at men and women the respective 75-25 ratios increased from about 2.3 to 2.7. This striking pattern is displayed in panel (b) of Figure 2, which reproduces Figure 11 in Guvenen et al. (2017). Although overall inequality in lifetime earnings changed very little, at the top of the distribution one does observe an increase in inequality: the 90-50 percentile ratio moved up from 2.3 to 2.7, while inequality at the bottom of the distribution fell, as indicated by the decreasing 50-10 percentile ratio, which dropped from 3.1 to 2.9.

Bowlus and Robin (2004) also study the evolution of lifetime earnings inequality over time. They note that lifetime income is dependent not only on an individual’s current position in the labour market, but also on the evolution of their wage and employment status. Therefore they develop a model incorporating wage mobility and employment transitions. For each individual in the sample they simulate a possible lifetime earnings trajectory, assuming that individuals are subject to the same distribution of shocks faced by older workers, and thereby estimate the distribution of lifetime earnings realizations. Their sample consists of white males aged 16-65 from the March CPS between 1978 and 1999. They use this sample to estimate the experience and education specific wage and employment transition probabilities. The key finding from this exercise is that lifetime earnings inequality is roughly 40% lower than annual earnings inequality. This implies that measures of cross-sectional inequality are significantly reduced when one factors in future employment and wages. This is partly due to the fact that the young, while exhibiting lower current wages, benefit disproportionately more from upward wage mobility relative to the old. It is also shown that inequality has been rising over time within education and experience groups, especially for university graduates. Both within and across group variation is important in explaining rising earnings inequality.

As mentioned above, using a narrow data sample from the NBER-TH survey, Lillard (1977) also examines lifetime dispersion of earnings and draws comparisons to cross-sectional earnings inequality. The sample for this classic study includes only males who volunteered for the U.S.
Air Force pilot, navigator and bombardier programs in the last half of 1943. Lillard explores the relative importance of schooling, measured ability and family background for both annual and lifetime earnings. His broad conclusions are that cross-sectional earnings inequality is about 50-80% higher than lifetime earnings inequality. Annual earnings inequality is high, both overall and by age group. These inequality measures are not sensitive to discounting, nor to the length of working life. Another finding of this pioneering work is that the contributions of schooling, ability, and background to variation in lifetime earnings are similar to their respective impacts on the variation of annual earnings within age groups; age itself is the most important factor in explaining annual earnings. These results, as well as those from the previously discussed articles, are echoed in a later study of Norwegian data by Aaberge et al. (2014). This study also shows that lifetime earnings are much more equally distributed than annual earnings. Lifecycle bias, which is assessed by comparing within-age earnings inequality to aggregate earnings inequality, is a key contributor to the discrepancy between annual and lifetime earnings.

**Expected Lifetime Income Inequality**

At least since the seminal work of Modigliani and Brumberg (1954) and Friedman (1957), it has been understood that expected lifetime earnings (as opposed to realized, and observed, lifetime earnings) are the main drivers of economic decisions and utility. This discrepancy between ex-ante (unobserved) magnitudes and ex-post realized outcomes is central to make sense of the choices and behaviours of economic agents. As Cunha et al. (2005) emphasize, differences in expected lifetime earnings reflect ex-ante heterogeneity in human capital, whereas un-forecasted differences in lifetime earnings reflect luck and uncertainty. Assuming concave utility functions and a lack of full insurance, greater inequality due to uncertainty reduces the welfare of all agents. However, greater inequality in expected (hence predictable) resources, holding the mean constant, would imply that the rich become better-off and poor become worse-off in welfare terms. Therefore, understanding how lifetime income inequality breaks down into its different components is crucial for understanding welfare.

A primary requirement to assess inequalities in human capital is to estimate expectations of lifetime earnings. That is, one needs to find ways to gather and aggregate all information available to individuals when making forward-looking judgements about their future stream of resources. These judgements about future earnings do not include variation due to the residual component of lifetime earnings, such as luck and other unforecastable factors, as these are, by definition, unrelated to a worker’s skills or potential to earn income. Two individuals can be identical in their education and earning abilities, yet they may realize very different earnings in
the future due to luck; however, despite these differences due to chance, one should consider
these individuals to have possessed the same ex-ante human capital when younger, and indeed
to have been equal in their opportunities. Two lines of research that use different approaches to
quantify inequality in the distribution of ex-ante human capital have been advanced by Huggett
and Kaplan (2016) and Abbott and Gallipoli (2019). The central questions in these lines of
inquiry relate to how agents discount future payoffs when forming expectations, given different
possible realizations of future earnings; and how much prior information (possibly unobservable
to researchers) agents possess about their future earnings. With respect to the former question,
recent advances in asset pricing theory suggest that agents employ state-dependent stochastic
discount factors, which can be estimated non-parametrically using consumption panel data.
With respect to the latter question, work by Cunha et al. (2005) suggests the use of early life
choices, such as the decision to complete education, as proxies of ex-ante information available
to agents. Credible answers to the question whether (and how much) realized lifetime earnings
inequality overstates ex-ante human capital inequality will be a critical, and likely contentious,
issue in the ongoing debate on human capital inequality.

**Permanent vs. Transitory Earnings Decompositions**

A popular approach to studying human capital inequality involves decomposing earnings into
underlying permanent and transitory components. While the transitory component reflects tem-
porary, unpredictable variation or measurement error in wages, the permanent component can
be thought of as a measure of human capital, akin to the flow value of the expected future earn-
ings discussed in the previous section. This decomposition turns out to be extremely useful
for two reasons: first, one can study trends in human capital inequality by estimating how the
variance of the permanent component of earnings has changed over time; second, the actual im-
plementation of this method can be carried out using fairly simple variance-covariance matrix
estimators. This approach does, however, require a stylized set of assumptions and cannot de-
 deliver point estimates of each individual’s permanent component value, nor does it suggest how
to discount across possible future realizations of earnings to generate human wealth estimates. That said, much has been learned about the nature of trends in human capital inequality through
such decompositions.

To fix ideas, denote an individual’s log earnings by \( y_{it} \), and assume they are the sum of
a transitory component \( \mu_{it} \) and permanent component \( \nu_{it} \): \( y_{it} = \mu_{it} + \nu_{it} \). The transitory
component is such that realizations \( \mu_{it} \) are independent of past and future realizations of the
same component, or more precisely \( E(\mu_{it}\mu_{it+\tau}) = 0 \) for \( \tau \neq 0 \). In contrast, the distribution of
realizations of the permanent component explicitly depends on past realizations. It is common for the permanent component to be modelled as either an $AR(1)$ or unit-root process; however, econometric advances also allow for non-linear dynamics, which may be a more appropriate assumption. Typically, these models can be identified with relatively short data panels, e.g. three year for the $AR(1)$ specification, but longer panels are needed for more reliable estimates. If longer longitudinal data sets are available, windows of data can be used to produce time-varying estimates of the variances of the key parameters of these models.

Textbook examples of this approach are Gottschalk et al. (1994) and Moffitt and Gottschalk (2012). The authors decompose the earnings residuals of prime age males in the Panel Study of Income Dynamics (PSID) from 1970 to 2004 into permanent and transitory components. As a baseline, they find that inequality measured by the variance of log-earnings residuals nearly doubled during this period, and between 51 and 69 percent of the increase is attributable to the permanent component of earnings. These findings are displayed in Figure 3 of Moffitt and Gottschalk (2012), which is reproduced in Figure 3 here. Because the permanent component better reflects human capital, as opposed to luck and other random variation, one can conclude that a large increase in human capital inequality occurred over the sample period, although only about half to two-third of the rise in earnings inequality is genuinely attributable to changes in the distribution of human capital. Their estimates for 2004 indicate that about 60% of earnings inequality is attributable to a permanent component.

Lochner and Shin (2014) also study the evolution of inequality in male earnings using PSID data from 1970 to 2008. They specifically focus on the returns to human capital and show that the pricing of unobserved skills has changed dramatically over time - returns to unobserved skills have increased in the 1970s and 80s before reversing back by the late 90s. Moreover, they observe that the variance in log earnings residuals remained stable until the 90s before rising. These discrepancies in trends point to changes in the variance of permanent shocks and transitory shocks. The authors also look at changes in returns at various points in the distribution of earnings (poorer vs richer individuals), and find that the increase in returns to unobserved skills is not detectable at the top of the earnings distribution.

Kopczuk et al. (2010) take earnings averaged over a five-year period to be a measure of the permanent component of earnings in the SSA data. Following this alternative approach they find that the permanent component is a much larger fraction of the variance of log-earnings, over 80% by the early 2000s. This approach also indicates that the entire rise in earnings inequality is due to a rise in the variance of the permanent component. It is not clear whether the differences in findings between this research and Moffitt and Gottschalk (2012) is due to the definition of permanent components, working with raw data rather than residuals, or the use of a different
While the previously mentioned articles have studied earnings (which depend both on wages and on labor supply), a closely related literature focused on the evolution of wages alone. The permanent component of one’s wage rate is also a measure of human capital, under a simple proportionality assumption. Using a permanent-transitory wage decomposition, estimated from PSID data, Heathcote et al. (2010c) study the evolution of the parameters of an AR(1) process describing the evolution of permanent wage factors. These authors find that both the permanent and transitory components of wages increased similarly from the mid 1960s to early 2000s, and that the permanent component is significantly larger than the transitory component. At the end of their sample period the transitory component accounts for only about 30% of the total variance of log-wage residuals.

Heathcote et al. (2010a) also work on residual wage inequality in the United States using the PSID data. Echoing the previously discussed literature, this research finds an increase in residual wage inequality since the mid 1960s, and finds that the rise is roughly 50% attributable to an increase in the permanent component of wages and 50% attributable to the transitory component. One key aspect discussed in this paper is the impact of alternative ways to identify the components of the wage process. Identification can be achieved using either a set of moments on the level of log-wages or a set of moments on first-differences of log-wages. Although the main findings just mentioned are not affected, the authors do note other important differences in the results obtained under the two identification strategies. This disagreement, they say, “indicates that the permanent-transitory model is misspecified.” This relates closely to the earlier comment suggesting that the literature has now directed itself towards studying non-linear dynamics.

Guvenen et al. (2015) study earnings dynamics using SSA data, and find that earnings shocks exhibit “strong negative skewness and extremely high kurtosis,” hence contradicting the log-normality assumption underlying the previously discussed literature. While these authors do not decompose earnings into permanent and transitory components in the same manner described above, they do study these components, and their relationships with inequality, using nonparametric methods.12 Their findings are very interesting. For high-earning individuals further increases in earnings tend to be transitory and decreases in earnings tend to be permanent. For low-earning individuals the opposite is true in that increases in earnings tend to be persistent and decreases tend to be transitory. It is not yet clear how this affects our understanding of the evolution of residual wage inequality; however, it does provide a further indication of why current (contemporaneous) earnings inequality is a poor measure of inequality in the stock of human capital.
International Perspective

Most of the studies discussed in this article refer to the United States. This is partly a reflection of the enormous amount of work that has been done using US data. Nonetheless, it is interesting to assess how the patterns of cross-sectional earnings inequality compare across different countries, and to pinpoint common trends and differences.

A good source of information about cross-sectional earnings dispersion in different countries is the special issue on inequality published by the Review of Economic Dynamics in 2010. This issue contains a selection of papers examining inequality patterns in nine countries (Canada, Germany, Italy, Mexico, Russia, Spain, Sweden, US and UK). A study of this detail and scale is rare, thus the findings are of some significance, as they allow one to compare aspects of inequality across the subset of countries considered. Most of the studies in this special issue are based on data for the period between the 1970s and the 2000s: this makes it possible to draw comparisons and assess similarities and discrepancies. Another complementary, and equally rich, source of information can be found in the work of Atkinson and Piketty (2007).

We begin by highlighting the most recurrent patterns. Most studies present similarities in the evolution of earnings and income inequality over the period considered: (1) in most countries one observes rising earnings inequality starting in the late 1980s, with the notable exceptions of Spain and Russia; (2) income inequality is significantly higher than consumption inequality, consistent with the notion that cross-sectional insurance and redistribution play a key role in all countries considered in the study; (3) earnings and income inequality grew remarkably during the 1990s in most countries, suggesting that common technological or institutional factors may have been key for inequality growth in that period; (4) residual inequality (wage and earnings dispersion after controlling for observable characteristics) played a key role in the rise of overall inequality, with transitory shocks accounting for a large share of this surge; (5) a commonality found across all countries considered is that the variance of permanent wage shocks is smaller than the variance of transitory wage shocks; (6) during recessions, inequality in earnings increased quite sharply everywhere, especially at the bottom of the distribution of earnings. This suggests a link between unemployment (or limited employment) and earnings of the poorest workers, who seem to carry most of the burden of regressions in terms of labor income.

Important differences also become apparent when comparing the inequality experience of different countries. For example, while labour supply plays a central role in shaping household level inequality dynamics in countries like the US and Canada, this is less true in European
countries like Italy, Spain and Germany. This discrepancy highlights the importance of heterogeneity in human capital utilization in different economies: whether labour supply changes at the household level impact the dynamics of aggregate inequality partly depends on the rate of female labour participation and on the distribution of household types (marital status, number of children, education of spouses). Some European countries exhibit a low level of female labour participation and, given the relatively inelastic male labor supply, this may weaken the pass-through from changes in hours worked to overall earnings dispersion. In addition, wage inequality in countries like Italy, Spain and Germany is measured to be much lower than in the US, which can make any heterogeneity in labor supply less salient for earnings dispersion.

The country that most closely mimics the patterns observed in the US appears to be Canada. As documented in the work of Brzozowski et al. (2010), the time paths for wage dispersion, hours dispersion and wage-hour correlation since 1975 are fairly close to those observed in the US. However, despite the growth in income inequality, disposable income and consumption inequality did not grow as much in Canada, suggesting that redistribution through taxation and benefits may have been effective in mitigating the effects of growing economic discrepancies. The relative mitigation effects seem, however, to have become weaker over time: during the 1980s there was a strong rise in before-tax income inequality in Canada, which was mostly absorbed by the tax and transfer system. However, after the 1990s, when the before-tax income inequality rose again, the tax and transfer system has been less effective in offsetting the rise in income inequality. This resulted in a more pronounced increase after-tax income inequality.

One interesting aspect that differentiates the US and Canada is the variation in the health dimension of human capital. Arguably, one of the key aspects determining the value of an individual’s human capital is the relative health enjoyed by that person. The value of human wealth crucially depends on the ability to generate income through labor supply, which becomes harder in poor health. In the United States a sharp decline in labour market, marriage, and health outcomes has been documented for relatively poorer white non-Hispanics. Case and Deaton (2017) find that, for birth cohorts after 1940, this demographic group, especially those with less than a four-year college degree, experienced a decline in real wages that is more pronounced with each successive cohort. This decline in real wages is accompanied by rising mortality from drugs and alcohol poisoning, suicide, higher risk of heavy drinking, chronic pain, labour force detachment, and declining marriage outcomes. These trends are not common to those with a bachelor’s degree; educated men have seen limited changes in health, mental health and marriage outcomes and have flat profiles for labour force participation, suicide and drug mortality. The decline in real wages is also not identical across education groups: controlling for age, real wages for those with degrees are on average 10% higher for the cohort born in 1980
relative to the cohort of 1940, while wages for those without a degree are 10% lower.

In related work, Milligan and Schirle (2018) present evidence that health inequality changes were not as extreme in Canada. Using a comprehensive administrative dataset of Canadian men and women spanning a half century, Milligan and Schirle examine the relationship between income and health, estimating a gap in life expectancy between the lowest and highest earners of about 11 percent (an eight years difference in lifespan for men). Comparing this gap to the one in the US, it is only about 3/4 as large as that estimated recently by Chetty et al. (2016). Crucially, these authors do not find the same reversal in survival rates for mid-life males that has been documented in US studies like Case and Deaton (2017). In contrast to the US experience, it appears that the evolution of the earnings-longevity gradient can be described as a fairly uniform shift in Canada, with equal improvements among both high and low earners. The fact that in the United States there is a growing mortality gap between the top and bottom of the income distribution, while in Canada the mortality gap remains fairly constant suggests the possibility that institutional factors may help mitigate the growth of human capital and earnings inequality. While no clear consensus exists on what exactly can account for such differences, some potential explanations could relate to differences in access to health-care and education, as well as the incidence of long-run stress and hardship associated with job uncertainty.18

The importance of income stabilizers is confirmed when examining the experience of Sweden, where earnings inequality increased in the early 1990s. As discussed in Domeij and Floden (2010) this growth is largely attributable to movements in and out of employment. These authors also find that earnings inequality in Sweden was mostly due to increasing residual earnings dispersion, triggered by increasing volatility of persistent shocks. Like in the case of Canada, inequality in disposable income (after taxes and transfers) increased much less than for raw incomes, indicating the presence of an effective welfare system. There is no evidence of an increasing trend in household level consumption inequality and, since persistent shocks are difficult to insure against, it is reasonable to conclude that the Swedish tax-and-transfer was able to absorb and mitigate the effects rising income inequality.19

Britain also exhibits patterns that are qualitatively similar to those observed in the US: income inequality in the UK rose dramatically during the 1980s and continued its growth over the 1990s, while consumption inequality did rise but at a slower pace. Interestingly, Blundell and Etheridge (2010) show evidence that the surge of inequality in Britain during the 1980s is attributable to the strong growth of the volatility of permanent shocks perturbing labour income. In turn, this suggests a role for structural transformation that affected both the cross-sectional distribution of, and returns to, human capital: as they point out, changing education differentials had a big role in this episode, especially for the growth of inequality over the 1980s and
early 1990s. The subsequent growth can instead be attributed to changes in transitory volatility, like in many other countries in this comparative study. This evidence is consistent with earlier results in Gosling et al. (2000).\(^{20}\)

One important caveat is that caution must be exercised when examining the drivers of inequality in different contexts, as local conditions and arrangements can affect both the distribution of human capital and its returns. One example of this is provided by the Mexican experience, as documented in Binelli and Attanasio (2010). Mexico’s market structure is quite different from that of its northern neighbours US and Canada. A distinctive feature of Mexico’s labor market is the relative size of its informal sector as well as one of the highest levels of measured income inequality in the world, exceeding that in the US. Inequality in Mexico rose most prominently during the 90s, when the ‘Peso crisis’ occurred and a possible explanation is that the share of workers in the informal sector responded to changes in aggregate conditions: given the unregulated labor market, unskilled workers are more vulnerable during economic downturns and when unemployment hits during an economic crisis, many workers see their wage negatively impacted as they accept informal employment. The cyclical sensitivity of the returns to labor has therefore a large effect on the poorer sections of the working population and affects inequality by stretching out the bottom end of the earnings distribution.

Another example of the central importance of local dynamics is Germany, a country with a unique historical background. In this case a major event - the reunification in 1990 - has shaped the recent evolution of inequality. Earnings dispersion was stable in the years before 1990, but after West and East Germany were rejoined, both wage and earnings inequality increased almost mechanically. Interestingly, the work by Fuchs-Schündeln et al. (2010) documents that the German public welfare system was rather effective in mitigating the impact of growing inequality, as disposable income and consumption inequality increased only modestly.

Two countries that bucked the overall trend of increasing inequality between the 1980s and early 2000s were Spain and Russia. In the case of Spain, Pijoan-Mas and Sánchez-Marcos (2010) show that inequality in individual net labour earnings and household net disposable income decreased significantly over the study period (1985-2000). A partial explanation can be found in the fact that the unemployment rate was extremely high at the beginning of the sample period, despite a continued economic expansion. Two key changes shaped the evolution of inequality in Spain after 1985: first, the tertiary education premium fell (in contrast to many other countries); second, the unemployment level fell from 24 percent in 1985 to 13 percent in 2000. These factors ultimately had the effect to reduce inequality in labour earnings in Spain, by compressing both the top and bottom ends of the distribution. The study also finds that the decrease in income inequality was driven mostly by changes in the distribution of permanent
components of earnings. This is in sharp contrast to the case of Russia, the other country where, in recent periods, recorded income inequality exhibited a downward trajectory. Analysis of Russian micro-data by Gorodnichenko et al. (2010) suggests that a moderation in the volatility of transitory shocks during a period of strong economic recovery was responsible for this pattern. Interestingly, expenditure and income inequality in Russia are not far apart, which potentially indicates a lack of cross-sectional insurance and an ineffective transfers system that is unable to mitigate fluctuations in disposable income and consumption.

While all these studies are useful to characterize the evolution of cross-sectional earnings inequality in a diverse set of countries, there is still relatively little work on the distribution of lifetime earnings and the present value of human capital in countries outside the US. As mentioned before, the lifetime notion of human wealth is appealing as it relates to extended flows of income associated with underlying productive skills. The lack of international evidence on lifetime human wealth inequality is possibly due to the fact that such studies require rich data and more complex analytical approaches. One notable non-US study on the relationship between current and lifetime income uses high quality administrative data from Norway (Aaberge et al., 2014). Specifically, this study examines longitudinal data between 1942 and 2006, considering cohorts born in the interval 1942-1944. In this way the authors are able to capture the full working life of the sample members between ages 23/25 and 62/64. The measure of lifetime income is computed using the approach of Haider and Solon (2006), and is consistent with studies in the American context focusing on the annuity value of the discounted sum of real income. Findings from this study are broadly consistent with those obtained for US data: they highlight the life-cycle bias in an individual’s earnings profile, and show that inequality measured using lifetime income is much lower than inequality measured using cross-sectional income.

**Ongoing Debate on Human Capital**

Heterogeneity in human capital is a key source of differences in economic well-being. This article provides a synopsis of the empirical approaches that have characterized the analysis of human capital inequality over the past few decades. In the process it overviews various methodological and measurement issues, and summarizes the existing evidence on the changing patterns in the distribution of returns to labor and human wealth across various countries. The focus is on specific aspects of human capital measurement and on the empirical evidence gathered by the large literature on these topics. It is impossible to summarize this extensive research entirely within such a limited space.
One key aspect on which this article remains silent is the analysis of the fundamental causes underlying the persisting, and often growing, disparities in human capital and earnings. Some of the potential causes examined in the literature relate to innate abilities, heterogeneity of early investments in human capital due to family background, as well as short term credit constraints that limit the ability to attend formal schooling, especially at the tertiary education level. Although much academic and policy focus has been on returns to schooling, several empirical studies have highlighted the importance of early environments in fostering cognitive and non-cognitive skills, which are key in determining the development of human capital and, eventually, economic success. This growing body of work has stressed the importance of childhood influences, especially at very young ages, on skill development, arguing that while institutional learning is an important aspect of skill development, it is not the only channel through which skills are developed, and not necessarily the most important (see Heckman, 2000; Aizer and Currie, 2014). The key message of this literature is that a child’s environment is a major predictor of future success, as disadvantages arise from lack of cognitive and non-cognitive stimulus early in life or even from beneficial environments at or before birth itself. These disadvantages are then compounded as the child moves through subsequent stages of development (see Heckman, 2004; Cunha and Heckman, 2007). While distinct, all these explanations find their motivation in the underlying lack of equal economic opportunities offered to young individuals.

The broad set of measures and strands of research discussed in this article all suggest that human capital inequality has grown since the late 1970s. However, since human capital is not directly observable, new results and empirical evidence can, and certainly will, be brought forward to further refine and advance our understanding of human capital heterogeneity. As stressed all along, earnings inequality is not the same as human capital inequality, and comparisons between annual and lifetime earnings inequality show how much transitory earnings inequality can overstate underlying human capital disparities. Lifetime earnings are themselves not a satisfactory proxy for human capital, as they encompass a large component of realized shocks that are unrelated to human capital. It is likely that the direction of future research in this field will lean towards data and methods that shed new light on individuals’ ex-ante, heterogeneous valuations of their human wealth. To understand the cross-sectional disparities in permanent income and human wealth it will be crucial to measure how much of realized lifetime income inequality is a predictable reflection of human capital at any point in time, as opposed to sheer luck or random unpredictable events. Answering this question has immediate and profound implications for tax and redistribution policies, as well as for the positive assessment of what drives individuals to vastly different long-term outcomes.
Notes

1. For example, a closely related literature on the relationship between consumption inequality and income inequality exists. Examples include Blundell et al. (2008), Aguiar and Bils (2015) and Blundell et al. (2016) among many others, as also summarized in the interesting book of Jappelli and Pistaferri (2017).

2. An obvious measurement issue is that the remuneration of human capital can be recorded as a payment for business activities and possibly be treated like capital income. Active business activity often takes the form of sole proprietorship of corporations organized as ‘limited liability’ legal entities that pay corporate income tax on annual taxable income. In such circumstances the owners can be active earners rather than passive rentiers, as pointed out by Smith et al. (2018) and Dyrda and Pugsley (2018). This complicates the measurement and taxation of human capital returns.

3. Non-trivial issues of measurement become apparent when looking at labor income payments, as pointed by Guvenen and Kaplan (2017). These authors study top income inequality, measured as the share of incomes accruing the individuals in the top percentiles of the income distribution: crucially they use a combination of total income data from the Internal Revenue Service (IRS) and labour income information from the Social Security Administration (SSA), starting from 1981. These two data sets allow them to isolate differences in top income inequality based on the definition of income (labour income vs total income) and to differentiate with respect to the unit of analysis (tax unit vs individual).

4. See also Quadrini and Ríos-Rull (2015)

5. Furthermore, Lemieux (2006a) shows that inequality that is not explained away by observables, such as education, increased very little.

6. Song et al. (2015) suggest that the firm, rather than the industry, is the key dimension.

7. Blundell et al. (2016) also estimate human wealth, i.e. expected lifetime earnings, but crucially they do no consider state-dependent stochastic discounting.

8. To learn about inequality across groups, for example, one must estimate a separate wage process for each partition. One such example is the work of Abbott et al. (2019) who estimate and simulate separate wage processes for six groups: men and women, each across three education levels.

9. See Arellano et al. (2017) or Guvenen et al. (2015) for recent research on nonlinear implementations.

10. What one can learn about the nature of earnings dynamics from longer auto-covariances is discussed in Meghir and Pistaferri (2006).

11. Residuals are defined as the component of log-earnings not predicted by education, race, an age polynomial, and interactions among these variables.

12. Theoretical work by Huggett et al. (2006) shows that some of these patterns can be replicated in a model of human capital dynamics.

13. See also Roine and Waldenström (2015) and Morelli et al. (2015).


15. For a discussion of possible explanations, see Lemieux (2006b).


17. For an informative comparison see Green and Sand (2015).

18. One aspect worth highlighting is that the concentration of wealth in Canada is less severe than that in the US: the top 5 percent of richest households holds 35 percent of wealth in Canada, which is similar to the share held by the top 1 percent of richest households in the US. A related literature examines the implications of changes in tax...
and transfers system for inequality, see Kaymak and Poschke (2016) and Hubmer et al. (2016).

Questions remain about the process of joint determination of human capital outcomes and redistributive policies. Abbott and Gallipoli (2017) explore the possibility that the shape of the distribution of human capital may itself affect optimal public policies and redistribution. They provide microdata evidence that this may in fact be the case for a set of developed countries.

See also Blundell et al. (2007).

References


——— (1964): *Human Capital*, NBER.


Figure 1: This is a reproduction of Figure 1 in Kopczuk et al. (2010). It displays the evolution of the earnings Gini index in SSA data from 1937 to 2004, both in aggregate and by gender.
Figure 2: This is a reproduction of Figure 11 in Guvenen et al. (2017). Panel (a) displays general patterns of earnings inequality measured by the standard deviation of log-earnings across genders and over time. Panel (b) displays 75-25 earnings ratios by gender and overall, showing how the pattern changes significantly if one considers only one gender or all workers. Panel (c) displays the evolution of the gender earnings gap, and panel (d) displays how within-gender inequality has risen while between-gender inequality has fallen.
Figure 3: This is a reproduction of Figure 3 in Moffitt and Gottschalk (2012). It displays the evolution of the total variance of log-earnings’ residuals from 1970 to 2004, as well as the evolution of both its permanent and transitory components.