

Koerselman, Kristian; Uusitalo, Roope

Working Paper

The Risk and Return of Human Capital Investments

IZA Discussion Papers, No. 7752

Provided in Cooperation with:

IZA – Institute of Labor Economics

Suggested Citation: Koerselman, Kristian; Uusitalo, Roope (2013) : The Risk and Return of Human Capital Investments, IZA Discussion Papers, No. 7752, Institute for the Study of Labor (IZA), Bonn

This Version is available at:

<https://hdl.handle.net/10419/89880>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

IZA DP No. 7752

The Risk and Return of Human Capital Investments

Kristian Koerselman
Roope Uusitalo

November 2013

The Risk and Return of Human Capital Investments

Kristian Koerselman

*Abo Akademi University,
HECER, SOFI and TCWR*

Roope Uusitalo

*HECER, University of Helsinki
and IZA*

Discussion Paper No. 7752
November 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

The Risk and Return of Human Capital Investments^{*}

Investing in human capital increases lifetime income, but these investments may involve substantial risk. In this paper we use a Finnish panel spanning 22 years to predict the mean, the variance and the skew of the present value of lifetime income, and to calculate certainty equivalent lifetime income at different levels of education. We find that university education is associated with about a half a million euro increase in discounted lifetime disposable income compared to vocational high school. Accounting for risk does little to change this picture. By contrast, vocational high school is associated with only moderately higher lifetime income compared to compulsory education, and the entire difference is due to differential nonemployment.

JEL Classification: C33, I24, J31

Keywords: lifetime income, risk, education, human capital

Corresponding author:

Kristian Koerselman
Department of Economics
School of Business and Economics
Åbo Akademi University
FI-20500 Turku
Finland
E-mail: koerselman@economistatwork.com

^{*} We thank the Network for Higher Education and Innovation Research at the University of Helsinki for financial support. We appreciate useful comments from Angela Djupsjöbacka, Markus Jäntti, Matthew Lindquist and Anders Stenberg as well as from participants at EALE, EEA/ESEM, ESPE, IWAE, the Educational Governance Workshop at the Norwegian University of Science and Technology, the 2nd Conference on Higher Education and Innovation Research at the University of Helsinki, and at seminars at Abo Akademi University, the Helsinki Center of Economic Research HECER, the Swedish Institute for Social Research SOFI, and at the Turku Center for Welfare Research TCWR.

1 Introduction

As any other investment, investment in human capital involves risk. At least to some extent, education produces skills that have market value in a limited set of occupations only. As the future demand for these skills is uncertain, students investing in education end up with a risky portfolio of specific skills. These risks are not insurable and cannot easily be diversified.

The value of an investment depends both on its expected return and on its risk. Hundreds of studies have demonstrated that higher education is associated with higher income. Much less is known about the risks associated with investments in human capital.

Education may affect both the transitory and the permanent component of earnings. Risks associated with year to year volatility of earnings can to some extent be smoothed through saving and borrowing, but risks related to the variation in lifetime earnings cannot. Since education is a long-term investment, with returns to investment distributed over the entire career, we argue that a natural measure of risk is the variability of lifetime income.

In this paper we measure the variability of lifetime income by estimating means, variances and skews of lifetime income by level of education. We use administrative data from Finnish registers where we can observe individual earnings over 22 years – a substantial fraction of careers though not entire lifetimes.

According to our estimates, higher education is generally associated with a higher mean, a higher variance and a higher skew of lifetime income. For a risk averse person, a higher variance decreases the value of education. The effect of skew is not discussed as often, but commonly used utility functions imply that skewness increases utility. Empirical evidence supports both variance aversion and skew affection (e.g. Golec & Tamarkin, 1998; Garrett & Sobel, 1999; Hartog, 2011). In this paper we use CRRA utility functions with varying assumptions on the degree of risk aversion to convert the moments of earnings distribution into certainty equivalent present values of lifetime earnings at different levels of education.

We find that risk-adjusted returns to education are comparable to unadjusted ones. According to our estimates, certainty equivalent lifetime earnings are about twice as large for university graduates as for vocational high school graduates, irrespective of whether we adjust for risk or not. When we account for taxes and transfers by using disposable income instead of earnings, this reduces the returns to education. Still, certainly equivalent lifetime disposable income is 60 to 80 percent higher for university graduates than for vocational high school graduates. On the other hand, the differences between vocational high school graduates and those with compulsory schooling only are much smaller, and mainly due to the difference in non-employment risk rather than to the differences in the earnings distribution among those who are employed.

Our paper is related to multiple strands of literature. The risk-return trade-off has been evaluated in a compensating differentials framework at least since King

(1974). Typically the approach involves adding measures of variance in a log earnings regression. Hartog (2011) provides a recent survey of this literature. Among the alternative approaches are Pereira & Martins (2002), who look at the risk and return of education using quantile regression on different countries, Christiansen, Joensen & Nielsen (2007), who compare the risks and returns across different fields of education, and Harmon, Hogan & Walker (2003), who model the return to education using a random coefficient model.

Our goals are similar to those of Brown, Fang & Gomes (2012), who also evaluate certainly equivalent gains from education after accounting for risk preferences, earnings volatility and progressive taxation. The key difference between our study and Brown, Fang & Gomes is that they use time separable utility function and evaluate the sum of utilities at each age while we directly evaluate the utility of lifetime income. Brown, Fang & Gomes assume that individuals are credit constrained while our approach essentially assumes that there are no restrictions on saving and borrowing. One of the benefits of our approach is that it allows us to account for years with zero or negative income (e.g. while at school).

Our paper is descriptive in nature and does not attempt to uncover the causal effects of education on the distribution of earnings. We make no effort to distinguish between risk and heterogeneity except for repeating the analysis for a smaller subsample where we can control for cognitive test scores and parents' level of education. Recent efforts of disentangling unobserved heterogeneity from uncertainty in residual earnings have resulted in conflicting conclusions (Cunha & Heckman, 2007; Chen, 2008; Mazza, van Ophem & Hartog, 2013). Without suitable instruments we have little to add to this debate.

Instead, we focus on reporting differences in the distribution of long-term incomes by the level of education both before tax and after accounting for taxes and transfers. In this sense our paper is related to a number of papers which compare annual to lifetime income distributions, e.g. Björklund (1993), Bönke, Corneo & Lüthen (2012) and Bhuller, Mogstad & Salvanes (2011).

In short, we add to the literature in a number of ways. First, we estimate the effect of education on the variance of income in a lifetime perspective where previous papers have typically used considerably shorter horizons. Second, in addition to the variance, we also account for the skew in the earnings distribution, separating upside risk from downside risk. Since our method is largely nonparametric, we do not have to log income measurements, and can therefore include zeros in the calculations of all three moments. Third, we account for employment risk and social insurance. Our approach makes it easy to deal with issues such as duration of education and early retirement as these can be directly observed from the data.

2 Data

We use the person file from the Finnish Linked Employee–Employer Data set (FLEED) compiled by Statistics Finland. It consists of a one third random

sample of individuals residing in Finland at some point between 1988 and 2009. Individuals are present in the data for each year that they are registered in the Finnish population register, and individual information can be linked across years using person identifiers.

FLEED contains data on the highest educational degree completed for each individual and year. These data are based on reports by educational institutions to Statistics Finland and contains information on the date and type of degree according to the Finnish Standard Classification of Education.

Key income variables in the data are earnings, taxable income, and taxes paid. Income information is based on tax records, and, as most transfers are taxable in Finland, they are also included in the data. We compute disposable income by subtracting taxes from taxable income. We use gross lifetime earnings and disposable lifetime income as the primary outcomes of interest in our analysis.

All monetary variables in FLEED are top-coded at a nominal level of EUR 200 000 throughout the years. Because censoring affects less than 0.06% of yearly observations, the effect of censoring on mean lifetime income is negligible. However, the higher moments are more sensitive to censoring. We alleviate this problem using data on taxes. Because tax information is censored at the same nominal level as income variables, we can impute incomes for the observations that have censored incomes but uncensored tax variables using average municipal tax rates and municipal taxes paid. Imputation of high incomes below the censoring threshold show that the accuracy of this method is more than satisfactory. Less than 0.002% of observations have censored municipal tax amounts, and for these we use the imputed amount at the municipal tax censoring threshold.

While municipal and church taxes are reported consistently in the data, state taxes are included only in some years. To ensure that tax treatment is comparable across years, we impute state taxes for all years by applying each year's tax schedule to that year's taxable earnings for all years. A comparison with the years for which state tax information is available shows that actual tax amounts are very close to those predicted by the tax schedule.

Since Finnish students make their first main educational decision after finishing compulsory schooling at age 16, we restrict the sample to those aged at least 16 at the end of the calendar year. Because different cohorts are subject to different pension systems, we also exclude those older than 64. Since old-age pensions are partially based on lifetime earnings however, relative differences in true lifetime disposable income can be thought to lie somewhere between relative differences in age 16–64 earnings and differences in age 16–64 disposable income.

We define the level of schooling as the highest level of schooling achieved at age 30, and we therefore also have to exclude individuals who never turn 30 within the sample period. Furthermore, since we can only observe the date and level of the highest degree in the years between 1988 and 2009, we have to exclude those who have received their highest degree after turning 30 but before entering the observation window because we do not know their highest degree at age 30.

We also discard immigrants by excluding persons who were citizens of a foreign country in any of the years that we have data on so that changes in immigrant composition over time does not affect our estimates.

The number of observations retained at each stage can be seen from Table 1. In total, we retain data 996471 individuals with on average about 17 yearly observations per individual.

Table 1: Sample selection.

	individuals	observations
between 16 and 64	1559554	22312681
excluding immigrants	1500141	21875181
schooling known	996471	16849188

Notes: the number of observations refers to the total number of yearly observations. The average number of years of data per individual in the final sample is thus about 17.

To make monetary variables comparable across time, we deflate all money amounts to 2009 prices level using the cost of living index from Statistics Finland. On top of that, we account for real earnings growth by deflating with an additional 2% per year, a figure that comes close to average real earnings growth both over the sample period and over longer time frames. In this way, we keep business cycle variation in the data. In calculating the moments of lifetime income, we assume that real income growth will be 2% in the future as well.

We compare both lifetime earnings and lifetime disposable income across levels of education. We also separate earnings risk from employment risk by examining separately the subsample of individuals that are either employed or in school. An individual is included in this subsample when he or she is either registered as a student during the last week of the year, or is registered as employed during the last week, has not been unemployed for more than two weeks during the year, and whose annual real earnings exceeded EUR 5000.

3 Schooling and income in Finland

Finnish children start school at age 7. All children attend comprehensive school for nine years. At age sixteen the students make the first important choices regarding their education. Currently, about 45% of students continue to vocational secondary education, which typically also takes three years to complete. After vocational school it is possible to continue at polytechnics or universities of applied sciences but many exit education after vocational school.

The other 55% of students enter three year general secondary programmes ending in a matriculation exam which provides eligibility for tertiary education. University students are accepted directly to programs leading to a Master's

level degree. University admission is competitive, with less than half of applicants being admitted. Universities do not charge tuition fees and students receive relatively generous student grants.

Dropping out is common at all levels. Currently about 15% of thirty-year-olds have no education after compulsory school. In most cases these individuals have entered vocational education but never graduated. Dropping out of university is also common. In the data university dropouts are usually coded as having a general secondary education as their highest degree.

While the Master's level university education and secondary vocational education have remained reasonably similar over a long time, the structure of education has changed much more at other levels. For example, universities awarded Bachelor degrees in short university programs during the 1970s. In the 80s, these disappeared but in the 90s they were re-created as vocational-oriented tertiary education at polytechnics. All these are coded as lower tertiary education degrees in the Standard Classification of Education but large changes in content complicate comparisons of cohorts born widely apart.

In this study we concentrate on comparisons between those with no post-compulsory education to those with vocational upper secondary education and to holders of Master's degrees. These education levels are consistently coded across cohorts, and jointly make up a large share of individuals in each cohort.

As a first glance at the relationship between schooling and long-term income differences, we plot the average long-term earnings by the level of schooling in Figures 1 and 2. In the figures we use data from the cohorts born between 1955 and 1964. For each individual in these cohorts we calculate 22-year average earnings using data from the years from 1988 to 2009, and we display the 10th, 25th, 50th, 75th and 90th percentiles of these averages. The figures also illustrate the size of the groups; the height of each box is proportional to the number of individuals at each education level.

As can be seen from the boxes' heights, compulsory education and vocational high school are the two largest education categories. From their position along the horizontal axis we can see that Master's degree holders earn substantially more than vocational high school graduates, but vocational high school graduates do not earn much more than those with compulsory school only. Variances and skews are larger for university graduates as well, while the variances and skews of the groups with vocational secondary and compulsory education are smaller and similar to each other.

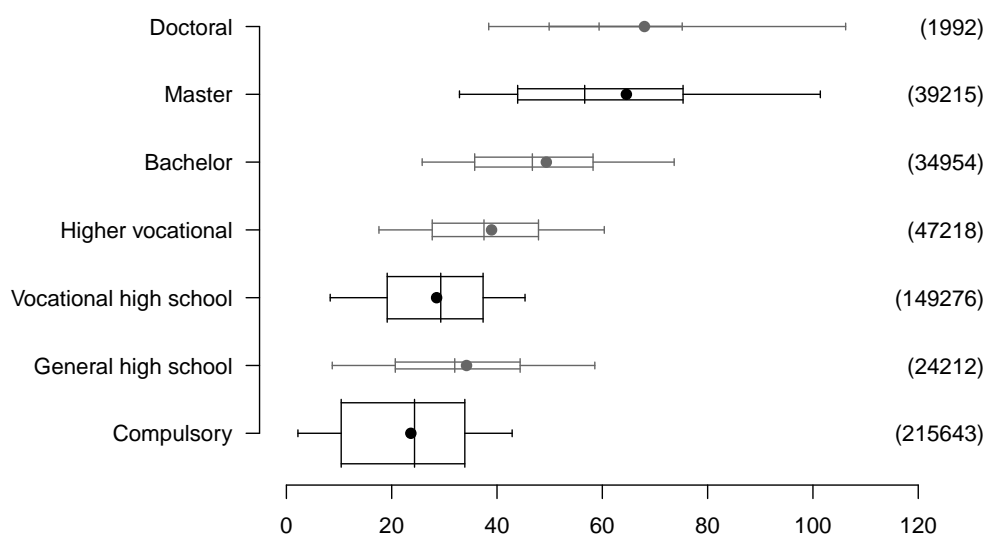


Figure 1: 10th, 25th, 50th, 75th and 90th percentiles of 22-year averages of annual earnings by education, men. Data cover cohorts born between 1955 and 1964. Earnings are observed from 1988 to 2009 and are expressed in thousands of 2009 EUR. Dots indicate means. The heights of the boxes are proportionate to the number of individuals in the sample, which also have been added within parentheses on the right hand side of the figure.

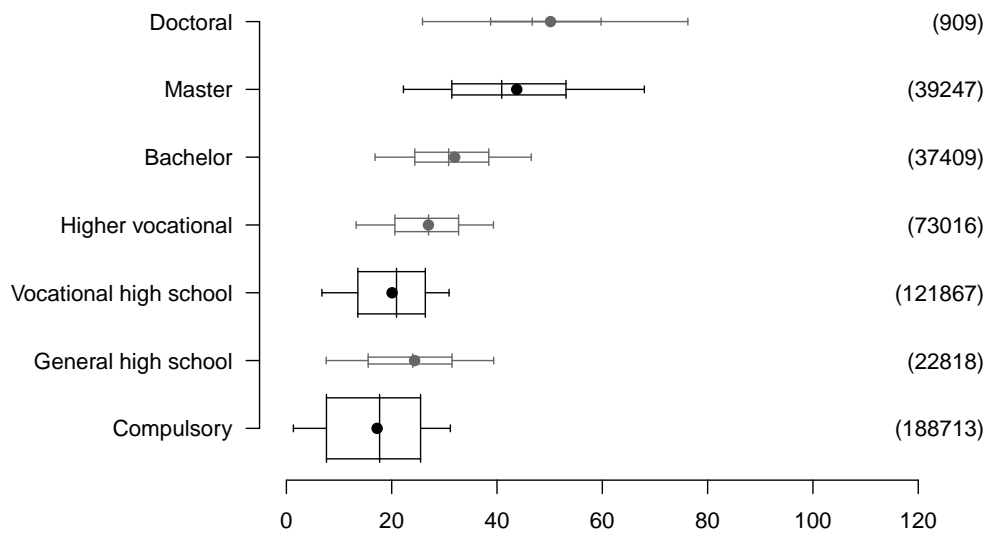


Figure 2: 10th, 25th, 50th, 75th and 90th percentiles of 22-year averages of annual earnings by education, women. See Figure 1 for details.

4 Methods

The relationship between education and lifetime income can be estimated based on a cross-section from a single year, and thus place relatively low requirements on the data. Calculating the variance of lifetime income is more demanding as it not only requires information about the variance of income at each age, but also about the covariance of income between different ages. Data sets spanning entire lifetimes are rare, but there are several data sets that span one or more decades. We demonstrate that such data can still be used to nonparametrically estimate the moments of lifetime income in a straightforward manner. We do this in a way that accounts for the risk of having no income and does not impose strong assumptions on the shape of age-income profile, length of schooling or working life, nor on the relationship between age and variance of income.

We make two simplifying assumptions: that cohort effects can be ignored and that time effects can be accounted for by simple trends as described in the data section. Under these assumptions, the mean lifetime income for any exogenously defined group can be nonparametrically estimated from cross-section or pooled panel data. Mean discounted lifetime income μ is simply the sum of the mean incomes w_t at each age t , discounted at rate r .

$$\mu = \mathbf{E}(Y) = \sum_{t=16}^{64} \frac{\bar{w}_t}{(1+r)^{t-16}}$$

In a similar way the variance of discounted lifetime income is the sum of the discounted elements of the variance-covariance matrix describing the covariances cov of income between ages t and j .

$$\mathbf{E}[(Y - \mu)^2] = \sum_{t=16}^{64} \sum_{j=16}^{64} \frac{cov(w_t, w_j)}{(1+r)^{t-16}(1+r)^{j-16}}$$

For all the cells of the covariance matrix to be known, we would need a panel spanning over the length of working life, but in that case we could calculate the variance of lifetime income directly. For shorter panels, only part of the covariance matrix will be observed. In general, when estimating a covariance matrix ranging over A age groups with a panel of length N , $(A - N + 1)(A - N)$ elements of the matrix will be unobserved. Since in our case $A = 49$ and $N = 22$, about 69% of the elements of the covariance matrix can be directly estimated. Fortunately these are also the elements that make the largest contribution to the variance of lifetime income because incomes at different ages are more correlated the closer together the ages are. Additionally, discounting reduces the weight of the missing elements.

We thus estimate nonparametrically covariances between residual incomes at age t and age $t + 1$ through $t + N - 1$, where N is the length of our panel. We impute the missing covariances $cov(w_t, w_j)$, $j > t + N - 1$ with the last covariance element that we could estimate, i.e. with the covariance of income at ages t

and $t + N - 1$. Under the assumption that covariances decrease monotonically from that point on, this will provide us with an upper bound of the missing covariances. As a robustness check, we repeat our analysis with imputed zeroes and with predictions from an AR(1) model instead. Because even at the upper bound only a small proportion of the total covariance is contained in the missing cells, these adjustments do not make a large difference in the results.

This method is easily extendible to higher moments, though the proportion of missing cells increases each time. For example, the skew of lifetime income is given by the sum of the discounted elements of the skew coskew tensor.

$$\mathbf{E}[(Y - \mu)^3] = \sum_{t=16}^{64} \sum_{j=16}^{64} \sum_{k=16}^{64} \frac{\text{coskew}(w_t, w_j, w_k)}{(1+r)^{t-16}(1+r)^{j-16}}$$

Because coskews drop off quickly away from the main diagonal, we make no attempts at imputing missing coskews, but simply set them to zero. This part of the analysis is also computationally intensive, since the skew coskew tensor has 49^3 cells for a working life of 49 years.

After having obtained the mean, variance and skew of lifetime income, we can enter them into an arbitrary indirect utility function U in order to compress them into a single metric. We use the constant relative risk aversion family of utility functions, which is given by

$$\begin{aligned} \frac{1}{1-\rho} Y^{1-\rho}, & \quad \rho \neq 1 \\ \ln(Y), & \quad \rho = 1. \end{aligned}$$

where Y is lifetime income, and ρ the coefficient of risk aversion. Because we cannot observe the actual distribution of lifetime incomes, but have estimates of its moments, we follow Hartog (2011), and instead use a third-order Taylor approximation of the utility function, into which we can substitute the first three moments of lifetime income.

$$U(Y) \approx U(\mu) + \frac{1}{2}(Y - \mu)^2 U''(\mu) + \frac{1}{6}(Y - \mu)^3 U'''(\mu)$$

With utility levels in hand, we can compute the certainty equivalent level of lifetime income (CE), which is the certain level of lifetime income which would yield the same expected utility as a risky draw from the estimated distribution of lifetime income

$$Y^{CE} = U^{-1}[\mathbf{E}(U[Y])].$$

We calculate the moments of lifetime income separately for men and women for three levels of education (compulsory, vocational high school, university). We also present separate estimates for earnings and disposable income and separate estimates for the employed only. We then calculate CEs for individuals with different levels of risk aversion using both a second and a third order Taylor approximation.

4.1 Standard errors

Since our estimators are essentially weighted sums of other estimators, it is technically possible to calculate standard errors analytically. This would however be extremely computationally intensive. Not even a conventional bootstrap lies within the limits of computing power available to us since it would for example require repeatedly calculating hundreds of thousands of coskews on millions of observations to estimate the standard error of the skew of lifetime income.

Instead, we follow Politis, Romano & Wolf (1999), and repeatedly draw without replacement 200 subsamples of 10 000 individuals from each education \times gender group of our original sample. Next, we estimate the standard errors of the subsample estimators from the distribution of the 200 subsample estimates of each statistic. Because the estimators converge in sample size n at rate $\tau_n = n^{-0.5}$, we then multiply the subsample standard errors by $b_s^{0.5} \cdot b_f^{-0.5}$, where b_s is the subsample size and b_f is the sample size for each education \times gender group in the full sample. This gives us the standard errors for the full sample estimators.

We also report standard errors for the differences in certainty equivalents between levels of education. Since the estimators of the certainty equivalents are independent by construction, these standard errors are simply given by

$$\sqrt{\hat{se}_h^2 + \hat{se}_l^2},$$

where \hat{se}_h^2 and \hat{se}_l^2 are the estimated standard errors of the certainty equivalent lifetime incomes of the higher and lower level of education respectively.

5 Results

5.1 Moments and certainty equivalent income

In Figure 3 we plot estimated age-earnings profiles. These earnings are inflated to 2009 price and real wage levels, but are otherwise simple arithmetic averages of earnings by age and education. Note that zero-earnings observations have been retained in the data used for calculating these averages.

From the figures we can confirm the conclusion from earlier Figures 1 and 2 that the university educated earn substantially more than those at the two other education levels, but that the differences between workers with vocational and workers with only compulsory education are not very large. Earnings for those with compulsory school only is initially slightly higher than for vocational school graduates, but they are quickly overtaken by vocational school graduates. University graduates overtake vocational school graduates after a few more years.

Figure 3 also shows that earnings are substantially higher than zero at ages when most individuals are still at school. This suggests that forgone earnings

while at school are not very large, making education look less costly. After graduation the earnings of university graduates increase rapidly. The growth of earnings among university graduates is much faster in both absolute and relative terms than at lower levels of education. The downward sloping part of the curves indicates early retirement, and the low levels of earnings at age 64 illustrate that we are not missing an overly large proportion of lifetime earnings by ending the observation period at age 64.

Figure 4 shows the estimated covariances of earnings between different ages for university educated men. On the diagonal we find the earnings variances at each age. The variance is the highest at high ages when some men are retired while others still have high earnings. The volume under the lines represents the (undiscounted) variance of lifetime earnings. As can be seen from the figure, the missing covariances constitute only a small proportion of the total variance of lifetime earnings.

The coskews with age 40 earnings of university educated men have been plotted in Figure 5. Again, these are based on 2009 wage and price levels, and are thus not taking into account either discounting or real wage growth. The shown plane is one of 49 possible cross-sections of the skew coskew tensor. As can be seen from the figure, the coskews drop off quickly. We therefore feel confident in imputing the missing coskews with zeroes.

Table 2 shows the estimated moments of lifetime income. Looking at the results for the entire sample in the top two panels of the table, we see again that mean lifetime income is much higher for university graduates than for the other two categories. The coefficient of variation follows a U-shaped pattern, being lowest for vocational high school graduates. The skews are clearly higher for the university educated than for the other categories. The moments of lifetime income are predictably smaller after taxes and transfers. Finally, all moments are smaller for women than for men, but overall patterns are similar across genders.

In the bottom two panels, we see the same measures calculated on a sample only including observations where the individual is either in school or working full time. Compared to the full sample, the coefficient of variation of lifetime income decreases and the relative skew increases quite substantially for the lowest two levels of education. More strikingly, individuals without any education beyond compulsory school have higher mean discounted lifetime incomes than those graduated from vocational high school once we condition on employment.

In Table 3, we combine the moments of lifetime income in different ways to arrive at certainty equivalent lifetime incomes. When assuming a coefficient of relative risk aversion ρ equal to 0, we are effectively calculating mean discounted lifetime income. When increasing the coefficient of relative risk aversion, the variances and skews of lifetime income gain in weight. As we saw before, university education is associated with substantially higher lifetime income than vocational high school. The difference is about one million euros for male earnings and about 700 thousand euros for female earnings. After taxes and transfers, these numbers are somewhat smaller, but still sizable at roughly 600 thousand euros and 400 thousand euros respectively.

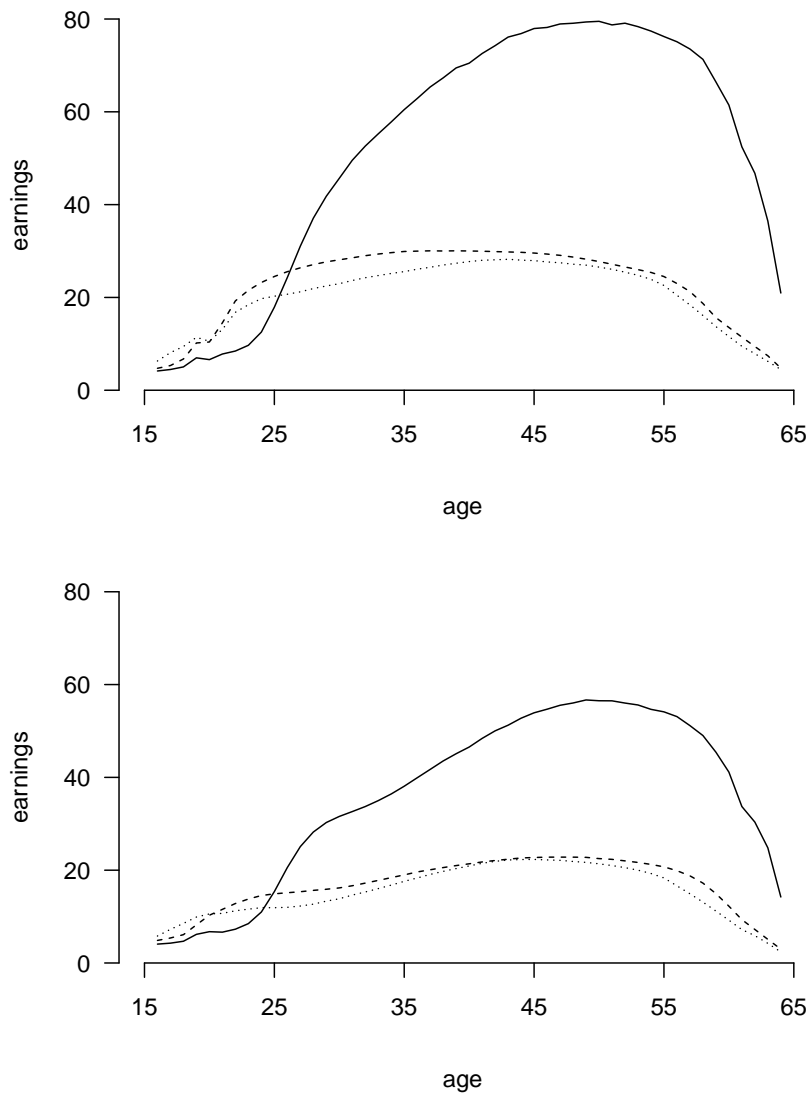


Figure 3: Cross-sectional earnings profiles in '000 EUR for university education (solid line), vocational high school (dashed line) and compulsory education only (dotted line): men (top panel) and women (bottom panel).

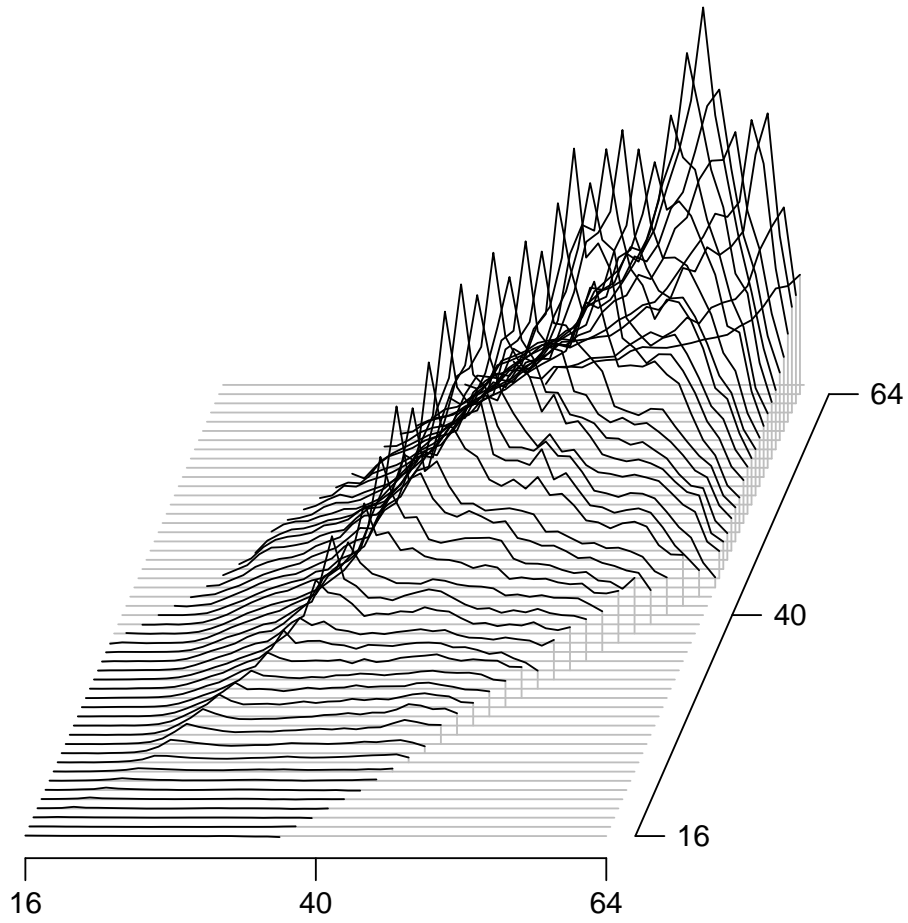


Figure 4: Covariances of residual earnings for university educated men. The volume under the lines represents the variance of lifetime earnings.

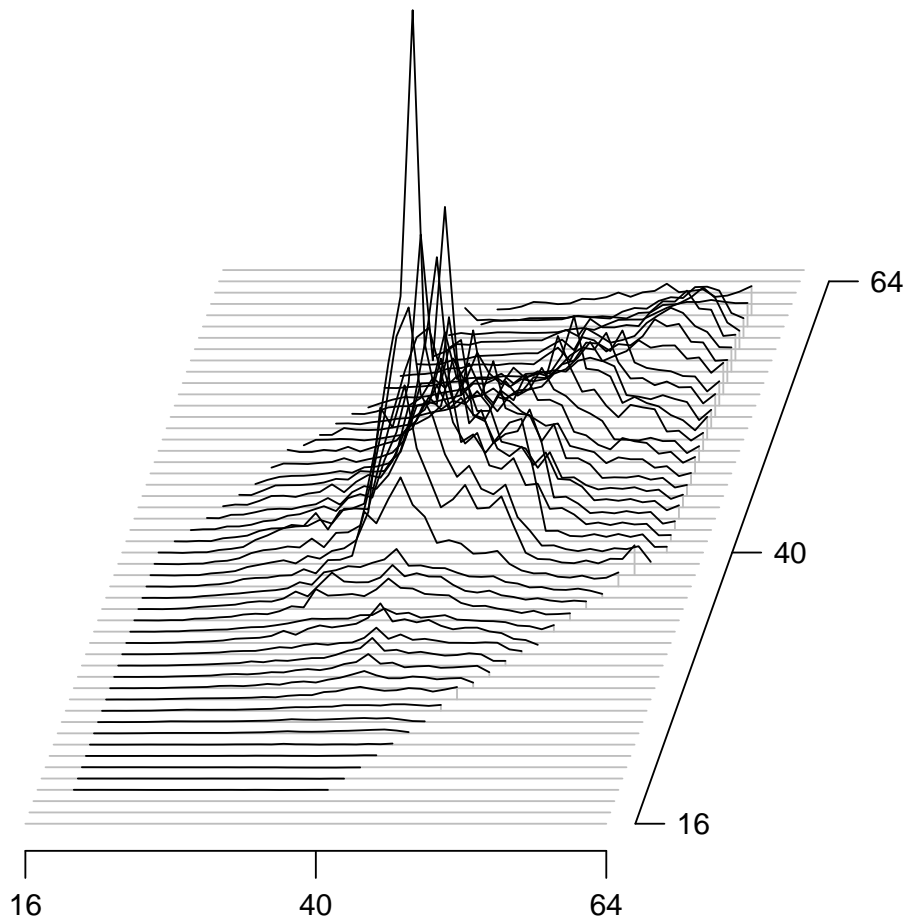


Figure 5: Coskews of residual earnings for university educated men. Shown is the cross-section of the skew coskew tensor at age 40. The joint volumes under the plots of the 49 possible cross-sections represent the skew of lifetime earnings.

One might think that the differences in mean lifetime incomes between university and vocational high school graduates are in fact compensating differentials for the higher variability of lifetime income for university graduates. After all, the coefficient of variation of lifetime income is higher for university graduates than for vocational high school graduates. However, when we look at the differences in CEs for higher levels of risk aversion, we see that they remain largely unchanged. This is because while variance enters the utility function negatively, skew enters it positively. Another way of putting this is to say that while university education is associated with higher levels of risk, this additional risk represents upside, not downside risk. The higher variability is thus not detrimental to expected utility.

We have also calculated CEs based on means and variances only using a second order Taylor approximation of the utility function. These estimates can be found in Table 4. When we omit the skew, the differences between university and vocational high school graduates are indeed lower at higher levels of risk aversion, though still substantial. This illustrates how important it is to take into account both the variance and the skew of income when evaluating income risk.

When we turn to the differences in CEs between vocational high school graduates and individuals with compulsory school only, we see that though the differences in CEs are small, they too increase in the coefficient of relative risk aversion. This is an indication of the more attractive risk profiles that vocational high school graduates have. The differences in CEs conditional on employment turn negative. This suggests that the higher CEs for vocational high school graduates is due to differential employment. An interesting observation is that while taxes and transfers are a net working life loss for most individuals, the mean disposable lifetime income of women with compulsory education only is higher than their mean lifetime earnings.

Table 2: Moments of discounted lifetime income.

men, entire sample									
	Compulsory			Vocational HS			University		
	mean	CV	skew	mean	CV	skew	mean	CV	skew
earnings	798 (2)	0.55 (0.00)	0.50 (0.05)	886 (2)	0.46 (0.00)	0.52 (0.11)	1914 (8)	0.52 (0.01)	2.59 (0.18)
disposable income	714 (1)	0.33 (0.00)	0.55 (0.07)	756 (1)	0.29 (0.00)	0.44 (0.12)	1311 (4)	0.39 (0.00)	2.48 (0.16)
women, entire sample									
	Compulsory			Vocational HS			University		
	mean	CV	skew	mean	CV	skew	mean	CV	skew
earnings	586 (1)	0.53 (0.00)	0.24 (0.03)	637 (2)	0.42 (0.00)	0.14 (0.02)	1327 (6)	0.45 (0.00)	1.61 (0.19)
disposable income	598 (1)	0.28 (0.00)	0.24 (0.02)	614 (1)	0.24 (0.00)	0.15 (0.02)	1005 (3)	0.31 (0.00)	1.59 (0.17)
men, employed or in school									
	Compulsory			Vocational HS			University		
	mean	CV	skew	mean	CV	skew	mean	CV	skew
earnings	1205 (2)	0.29 (0.00)	1.68 (0.18)	1193 (4)	0.28 (0.00)	1.53 (0.26)	2120 (9)	0.47 (0.01)	2.78 (0.18)
disposable income	878 (1)	0.23 (0.00)	1.60 (0.21)	884 (2)	0.22 (0.00)	1.06 (0.25)	1383 (5)	0.37 (0.00)	2.66 (0.16)
women, employed or in school									
	Compulsory			Vocational HS			University		
	mean	CV	skew	mean	CV	skew	mean	CV	skew
earnings	934 (1)	0.24 (0.00)	1.05 (0.12)	893 (2)	0.22 (0.00)	0.86 (0.07)	1511 (7)	0.39 (0.01)	2.08 (0.20)
disposable income	738 (1)	0.19 (0.00)	0.62 (0.05)	718 (1)	0.17 (0.00)	0.45 (0.04)	1076 (4)	0.29 (0.00)	2.12 (0.20)

Notes: Means are in '000 EUR. The coefficient of variation CV is defined as the standard deviation divided by the mean. The measure of skew reported is the third moment about the mean divided by the third power of the standard deviation.

Table 3: Certainty equivalent lifetime income, third order Taylor approximation.

men, entire sample									
relative risk aversion ρ	Compulsory			Vocational HS			University		
	0	1	2	0	1	2	0	1	2
earnings	798 (2)	705 (2)	654 (4)	886 (2)	811 (3)	764 (6)	1914 (8)	1889 (14)	2121 (37)
difference				88 (3)	106 (4)	110 (7)	1029 (9)	1078 (14)	1357 (37)
disposable income	714 (1)	680 (1)	655 (1)	756 (1)	727 (1)	704 (2)	1311 (4)	1276 (5)	1305 (9)
difference				42 (2)	47 (2)	49 (2)	555 (4)	549 (5)	601 (9)
women, entire sample									
relative risk aversion ρ	Compulsory			Vocational HS			University		
	0	1	2	0	1	2	0	1	2
earnings	586 (1)	515 (2)	471 (2)	637 (2)	586 (2)	547 (2)	1327 (6)	1260 (7)	1258 (14)
difference				51 (2)	70 (2)	76 (3)	690 (6)	674 (7)	711 (14)
disposable income	598 (1)	576 (1)	557 (1)	614 (1)	596 (1)	581 (1)	1005 (3)	973 (3)	958 (4)
difference				15 (1)	21 (1)	25 (1)	391 (3)	376 (3)	377 (4)
men, employed or in school									
relative risk aversion ρ	Compulsory			Vocational HS			University		
	0	1	2	0	1	2	0	1	2
earnings	1205 (2)	1170 (2)	1154 (4)	1193 (4)	1160 (3)	1142 (5)	2120 (9)	2092 (13)	2285 (31)
difference				-12 (4)	-10 (4)	-12 (7)	927 (10)	932 (13)	1143 (32)
disposable income	878 (1)	860 (1)	849 (2)	884 (2)	866 (2)	853 (2)	1383 (5)	1351 (5)	1382 (9)
difference				6 (2)	6 (2)	4 (3)	499 (5)	485 (6)	529 (9)
women, employed or in school									
relative risk aversion ρ	Compulsory			Vocational HS			University		
	0	1	2	0	1	2	0	1	2
earnings	934 (1)	913 (2)	896 (2)	893 (2)	875 (2)	860 (2)	1511 (7)	1459 (7)	1469 (12)
difference				-41 (2)	-38 (2)	-36 (3)	618 (7)	584 (7)	609 (12)
disposable income	738 (1)	726 (1)	716 (1)	718 (1)	709 (1)	700 (1)	1076 (4)	1050 (4)	1042 (5)
difference				-20 (1)	-18 (1)	-16 (2)	358 (4)	341 (4)	342 (5)

Notes: Values in '000 EUR, discounted to age 16 at $\delta = 0.03$. Utility is CRRA(ρ). Standard errors in parentheses.

Table 4: Certainty equivalent lifetime income, second order Taylor approximation.

men, entire sample									
	Compulsory			Vocational HS			University		
relative risk aversion ρ	0	1	2	0	1	2	0	1	2
earnings	798	685	612	886	798	733	1914	1669	1503
	(2)	(2)	(2)	(2)	(2)	(2)	(8)	(7)	(8)
difference				88	112	121	1029	872	770
				(3)	(3)	(3)	(9)	(7)	(8)
disposable income	714	676	643	756	725	697	1311	1215	1138
	(1)	(1)	(1)	(1)	(1)	(1)	(4)	(3)	(4)
difference				42	49	54	555	490	441
				(2)	(2)	(2)	(4)	(4)	(4)
women, entire sample									
	Compulsory			Vocational HS			University		
relative risk aversion ρ	0	1	2	0	1	2	0	1	2
earnings	586	509	457	637	584	542	1327	1199	1102
	(1)	(2)	(2)	(2)	(2)	(2)	(6)	(5)	(5)
difference				51	74	85	690	615	560
				(2)	(2)	(2)	(6)	(5)	(5)
disposable income	598	575	554	614	596	580	1005	957	916
	(1)	(1)	(1)	(1)	(1)	(1)	(3)	(2)	(2)
difference				15	21	26	391	361	336
				(1)	(1)	(1)	(3)	(3)	(3)
men, employed or in school									
	Compulsory			Vocational HS			University		
relative risk aversion ρ	0	1	2	0	1	2	0	1	2
earnings	1205	1154	1109	1193	1147	1107	2120	1895	1731
	(2)	(2)	(2)	(4)	(3)	(3)	(9)	(7)	(8)
difference				-12	-6	-2	927	747	624
				(4)	(4)	(4)	(10)	(8)	(8)
disposable income	878	855	833	884	863	844	1383	1291	1215
	(1)	(1)	(1)	(2)	(2)	(2)	(5)	(4)	(4)
difference				6	9	11	499	427	371
				(2)	(2)	(2)	(5)	(4)	(4)
women, employed or in school									
	Compulsory			Vocational HS			University		
relative risk aversion ρ	0	1	2	0	1	2	0	1	2
earnings	934	908	885	893	872	853	1511	1399	1309
	(1)	(1)	(2)	(2)	(2)	(2)	(7)	(5)	(5)
difference				-41	-36	-32	618	527	456
				(2)	(2)	(2)	(7)	(6)	(5)
disposable income	738	725	713	718	708	699	1076	1032	993
	(1)	(1)	(1)	(1)	(1)	(1)	(4)	(3)	(3)
difference				-20	-17	-15	358	324	295
				(1)	(1)	(2)	(4)	(3)	(3)

Notes: Values in '000 EUR, discounted to age 16 at $\delta = 0.03$. Utility is CRRA(ρ). Standard errors in parentheses.

5.2 Selection on observables

Though moments of lifetime income are both academically interesting and policy relevant of their own, it would also be good to have an indication of how much of the differences in moments are due to selection. This is a difficult problem even if suitable instruments for estimating causal effect of education on mean income was available. Unfortunately, we do not have reliable instruments for this sample.

A partial remedy is provided by test scores from verbal, math and logical reasoning tests administered to conscripts, which are available for a small number of male cohorts. For these cohorts, information on parents' education is available as well. The length of the panel is more limited in this subsample: earnings can be observed for the years 1995 through 2003.

We take the 1965 cohort, for which we observe highest education at age 30 as well as their earnings at ages 30 through 38. We sum these yearly earnings to end up with a proxy of lifetime earnings, and take the first three moments of the summed earnings by education.

We then repeat this exercise, but regress summed earnings on the three test scores and their squares as well as on dummy variables for each level of paternal and maternal education. From this regression, we take the squared and cubed residuals and regress them in turn on the same covariates. Because we cannot impute censored earnings in this data set, we instead censor residuals at the 99th percentile for each level of education.

We then predict mean summed earnings by education using the first regression, holding the covariates fixed at their mean levels in the population. From the second regression we predict the variance of summed earnings, and from the third the skew of summed earnings. We then recombine these predicted moments into certainty equivalent lifetime earnings like before, and scale them relative to the certainty equivalent lifetime earnings for vocational high school graduates.

The results can be seen from Table 5. Panel (a) shows the unconditional differences in certainty equivalent lifetime earnings reported earlier, but this time reported as the proportional premium over vocational education. Panel (b) shows the premia when we restrict our sample to the years 1995–2003. These are close to each other.

In panels (c) and (d), we show unconditional risk premia using the method described above; in panel (c) for the FLEED data set and in panel (d) for the army sample. These two sets of estimates are also close to each other but lower than in the full FLEED sample because this sample is observed at a younger mean age.

Panel (e) shows educational premia conditional on test scores and parental education. As expected, these are smaller, reflecting selection into education based on observable characteristics. Panel (f) shows how much of each premium is due to selection on observables. The mean difference between university and

vocational high school earnings is 22% smaller when holding covariates constant. It is 21% smaller when looking at certainty equivalent earnings using the CRRA(1) utility function, and 17% smaller using CRRA(2). A similar proportion of earnings differences between vocational secondary school graduates and individuals with compulsory school only can be explained away using observed covariates.

These results show that while the differences in lifetime earnings distributions documented in this paper are probably partially due to selection into education, the impact of selection seems to be about equally large with or without risk adjustments. The fact remains that income differences by education are large in Finland, and that adjusting for risk does very little to change this. At the very least, this calls into question the popular Finnish belief that education is a risky investment.

Table 5: Relative unconditional and conditional observed certainty equivalent lifetime earnings for three levels of risk aversion, proportional difference to vocational earnings. Men only.

ρ	0	1	2
<i>(a) FLEED baseline, unconditional</i>			
compulsory	-10%	-13%	-14%
university	+116%	+133%	+178%
<i>(b) FLEED 1995-2003, unconditional</i>			
compulsory	-14%	-20%	-23%
university	+119%	+132%	+160%
<i>(c) FLEED 1995-2003, 1965 cohort, unconditional</i>			
compulsory	-22%	-27%	-29%
university	+ 91%	+ 98%	+106%
<i>(d) army sample 1995-2003, 1965 cohort, unconditional</i>			
compulsory	-20%	-25%	-27%
university	+90%	+100%	+117%
<i>(e) army sample 1995-2003, 1965 cohort, conditional</i>			
compulsory	-15%	-20%	-21%
university	+70%	+79%	+97%
<i>(f) proportion of premium due to selection on observables = $(d - e)/d$</i>			
compulsory	0.25	0.20	0.22
university	0.22	0.21	0.17

6 Conclusions

Education is a lifetime investment that needs to be evaluated using data over the entire working life. Panels spanning working lives are however not usually available. Panels spanning ten to twenty years on the other hand do exist in many countries, in particular in countries where such data can be collected from administrative registers.

In this paper we show how moments of lifetime income can reliably be estimated from a shorter panel. Using a 22-year panel based on the Finnish registers, we find mean discounted lifetime earnings differences between university graduates and vocational high school graduates of up to about a million euros. These differences persist in certainty equivalent values after adjusting for the variance and skew of incomes. Moments of lifetime income are predictably smaller after taxes and transfers, but the major part of the income difference remains.

When we compare lifetime incomes of vocational secondary graduates to the incomes of those with compulsory school only, we find only small income differences. After controlling for employment, compulsory school graduates have higher discounted lifetime incomes. This suggests that vocational secondary education is mainly a hedge against nonemployment in Finland.

References

- Bhuller, M., Mogstad, M., & Salvanes, K. (2011). Life-cycle bias and the returns to schooling in current and lifetime earnings. IZA Discussion Paper, 5788.
- Björklund, A. (1993). A comparison between actual distributions of annual and lifetime income: Sweden 1951–89. *Review of Income and Wealth*, 39, 377–386.
- Bönke, T., Corneo, G., & Lüthen, H. (2012). Lifetime earnings inequality in germany. CEPR Discussion Paper, 8929.
- Brown, J., Fang, C., & Gomes, F. (2012). *Risk and Returns to Education*. Working Paper 18300 National Bureau of Economic Research.
- Chen, S. (2008). Estimating the variance of wages in the presence of selection and unobserved heterogeneity. *The Review of Economics and Statistics*, 90, 275–289.
- Christiansen, C., Joensen, J. S., & Nielsen, H. S. (2007). The risk-return trade-off in human capital investment. *Labour Economics*, 14, 971 – 986.
- Cunha, F., & Heckman, J. (2007). Identifying and estimating the distributions of ex post and ex ante returns to schooling. *Labour Economics*, 14, 870–893.
- Garrett, T., & Sobel, R. (1999). Gamblers favor skewness, not risk: Further evidence from united states’ lottery games. *Economics Letters*, 63, 85–90.
- Golec, J., & Tamarkin, M. (1998). Bettors love skewness, not risk, at the horse track. *Journal of political economy*, 106, 205–225.
- Harmon, C., Hogan, V., & Walker, I. (2003). Dispersion in the economic return to schooling. *Labour Economics*, 10, 205–214.
- Hartog, J. (2011). A Risk Augmented Mincer Earnings Equation? Taking Stock. In S. W. Polachek, & K. Tatsiramos (Eds.), *Research in Labor Economics*, vol. 33 (pp. 129–173). Emerald Group Publishing Limited.
- King, A. (1974). Occupational choice, risk aversion, and wealth. *Industrial and Labor Relations Review*, 27, 586–596.
- Mazza, J., van Ophem, H., & Hartog, J. (2013). Unobserved heterogeneity and risk in wage variance: Does more schooling reduce earnings risk? *Labour Economics*, (pp. –).
- Pereira, P., & Martins, P. (2002). Is there a return–risk link in education? *Economics Letters*, 75, 31–37.
- Politis, D., Romano, J., & Wolf, M. (1999). *Subsampling*. Springer Verlag. New York.