IZA

DISCUSSION PAPER SERIES

IZA DP No. 4069

Do Students Expect Compensation for Wage Risk?

Juerg Schweri
Joop Hartog
Stefan C. Wolter

March 2009

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Do Students Expect Compensation for Wage Risk?

Juerg Schweri
Swiss Federal Institute for Vocational Education and Training

Joop Hartog
University of Amsterdam and IZA

Stefan C. Wolter
University of Berne and IZA

Discussion Paper No. 4069
March 2009

IZA

P.O. Box 7240
53072 Bonn
Germany

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

IZA Discussion Paper No. 4069
March 2009

ABSTRACT

Do Students Expect Compensation for Wage Risk?*

We use a unique data set about the wage distribution that Swiss students expect for themselves ex ante, deriving parametric and non-parametric measures to capture expected wage risk. These wage risk measures are unfettered by heterogeneity which handicapped the use of actual market wage dispersion as risk measure in earlier studies. Students in our sample anticipate that the market provides compensation for risk, as has been established with Risk Augmented Mincer earnings equations estimated on market data: higher wage risk for educational groups is associated with higher mean wages. With observations on risk as expected by students we find compensation at similar elasticities as observed in market data. The results are robust to different specifications and estimation models.

JEL Classification: D8, I2, J2, J3

Keywords: wage, expectations, wage risk, risk compensation, skewness

Corresponding author:

Joop Hartog
University of Amsterdam
Department of Economics
Roeterstraat 11
1018 WB Amsterdam
The Netherlands
E-mail: j.hartog@uva.nl

* The authors thank Bernhard A. Weber for data assistance and assistance with the computer based survey, and Marc Fuhrer for technical assistance. For helpful discussions we thank Adam Booij, conference participants at EALE, at the Verein für Socialpolitik and at seminars at the Universities of Amsterdam and Zurich. Part of the work on this paper was done while Hartog was a Wertheim Fellow at Harvard.

1 Introduction

What do potential students know about the financial consequences when they decide on their education? What about the uncertainty about their potential wages? Are they aware of it, is there variation between individuals? Do they expect compensation for the risk in future earnings just as the stock market is known to present a trade-off between returns and risk? Individuals' choices on education are inevitably made under conditions of uncertainty, and thus expectations on consequences are essential input in the decision process. Yet about expectations we know very little. Only a handful of studies use information from questionnaires asking for individuals' expected mean or median wages that would ensue after some specified schooling (e.g., Kodde 1986; Betts 1996; Wolter and Zbinden 2001; Nicholson 2002; Webbink and Hartog 2004; Brunello et al. 2004). Dominitz and Manski (1996) extended this approach with eliciting information on the uncertainty of expected incomes.[1]

Although it is an obvious way to start research on the nature, relevance and impact of individuals' expectations, Friedman's methodology of instrumentalism seems to have precluded widespread adoption of the obvious approach (Friedman 1953).[2] The common method in empirical work is to treat expectations as unobserved variables and model the way individuals are supposed to extract the relevant input from information that is available to the researcher (including the possible filtering effect of self-selection).

In this paper, we will use the direct approach and analyse data obtained from students on the expected consequences of schooling choices. As one of the advantages of this approach we may note that it solves the problem of the counterfactual in a natural way: benefits from schooling options that are not chosen are not measured indirectly by construction, but simply by asking what individuals would expect.

Standard human capital theory assumes that students take into account the expected (discounted) lifetime income of different educational and occupational pathways when deciding about their education and occupation. However, if students are risk averse, the decision will not only be based on the expected value of lifetime income profiles, but also on the risk that is associated with each pathway. While this has been acknowledged at least since the early

[1] Zafar (2007) used probabilistic questions on abilities and work preferences, and used this information to explain choice of college major.

[2] Manski (2004) assesses the status of direct measurement of individuals' expectations and discusses examples of relevant use.

seventies (see Weiss 1972, and Levhari and Weiss 1974), the empirical literature dealing with wage risk is still scant, although interest is growing.[3] In our analysis of expectations, we will pay particular attention to the earnings risk that is associated with schooling choices. We will also investigate to what extent individuals' expectations reflect compensation for earnings risk. A small literature, exemplified by Hartog (2007), claims that wage levels for given education and occupation include a compensation for the risk emanating from imperfect predictability of wages at the time that individuals have to make their education-occupation investment choices. Wages are indeed higher in occupations-educations where wage variance is higher and lower in occupations-educations where skewness is higher: people dislike risk but appreciate favourable odds of very high wages relative to very low wages ("skewness affection").
A main criticism addressed at this approach is that the variance in the distribution of wages for a given education may not be a valid proxy for the ex ante wage risk faced by the agents. The observed distribution will also reflect heterogeneity and may be twisted and truncated by selectivity as individuals act on their private information. To tackle these problems, we will turn directly to the individuals themselves and ask about the wage distributions they expect under different age-education scenarios. By definition, students' individual wage risk expectations are not distorted by heterogeneity, as might be the case with aggregate wage risk measures derived from market data. Applying the methodology developed by Dominitz and Manski (1996) to Swiss students, we construct ex ante measures of wage risk and skewness.
Expectations data are particularly suited for the question at hand, since market compensation for wage risk has to be imposed by supply reactions: without sufficient compensation, students will not enter a specific education. For wage risk compensation to materialise students should be aware of risk and evaluate wages in view of that risk; in other words, risk compensation should be found in expectation data, in much the same way as stock market investors anticipate higher returns for more risky assets. In this paper we contribute to the small empirical literature on uncertainty in schooling decisions by testing awareness of compensation for earnings risk. We find that students expect significant wage risks and that higher expected wage risk is associated with higher expected mean wages.

The paper is structured as follows: Section 2 explains how we measure expected wage dis-

[3] For a survey of the literature and the different approaches used, see also the Special Issue of Labour Economics, December 2007, on Education and Risk.

tribution parameters and presents the data. Section 3 tests the hypothesis of expected risk compensation applying earnings regressions to expectation data and Section 4 concludes.

2 Calculating risk and skewness from wage expectation data

2.1 How to elicit wage expectations?

Dominitz and Manski (1996) did pioneer work in eliciting wage expectations of students. While there exists a literature using mean or median wage expectations, Dominitz and Manski asked students not only to state their expected median wage under different, specified age-education scenarios, but asked for additional information on the expected wage distribution. With this information, they were able to fit log-normal wage distributions for every student and every scenario. Their sample consisted of 110 US students who were surveyed via computer-assisted self-administered interviews.

Dominitz and Manski (1996) mainly focus on discussing the methodology of the survey and on providing evidence that the expectation data is informative and reliable. This is done by considering the internal consistency of the answers, the prevalence of response patterns, and the comments made by respondents in a debriefing session. Dominitz and Manski conclude from their analyses that "respondents are willing and able to respond meaningfully to questions eliciting their earnings expectations in probabilistic form" (Dominitz and Manski 1996, p.1).

The data used in the present study was gathered with computer-assisted interviews similar to those used by Dominitz and Manski. The survey was administered to four successive cohorts, 1998 to 2001, of students in the Economics Department of the University of Applied Sciences in Berne. 252 students were surveyed in their first semester.[4]

First, students were asked to give their expected median wage for a specified age/education scenario. Then wage distribution information for this scenario was gathered by defining wage values 20 percent below and 20 percent above the median stated by each respondent. These values were rounded to multiples of 500 Swiss francs, to prevent confronting students with odd and distracting values. The students had to state their perceived probability that they will earn at most 80 percent of the median and at least 120 percent of the median respectively. Thus, one has three points of the individual expected wage distribution for which the wage

[4] Wolter (2000) describes the software used for the survey in more detail. We used the same software here, but with a new sample of students. The exact phrasing of the questions can be found in Appendix B.

value and the position in the probability distribution are known. This procedure has been used for a total of ten different scenarios which vary age and education for own expectations and for perceived actual market wages.

The computer software provided the respondents with information needed to understand the probability questions (e.g., definition of the median) and checked the answers in real-time for missing or inconsistent values. The software also offered interactive help in case of errors. Furthermore, information about personal characteristics was gathered, such as gender, age, parents' education, parents' social class and grades in secondary school. Finally, students were asked to express their agreement with different normative statements.

Our sample has some major advantages. First, the sample size (although limited) is more than double that of comparable studies in the past. Second, the sampling was restricted to a well defined and homogenous group of students, limiting the risk that inter-individual differences would be driven too much by institutional or individual background variables. Third, there are no problems related to selectivity of participation. All students in all classes at the University of Applied Sciences participated in the survey (during class hours). Last but not least, the data is of higher quality than the data from written surveys. Item non-response or implausible answers are almost inexistent, thanks to the real-time plausibility checks of the software. Thus, hardly any observations drop out of the estimations. This rules out another important potential source for selection bias.

2.2 Operationalization of risk and skewness

As described in section 2.1, the students were asked to give information on probabilities for each of the intervals defined by the values corresponding to 80 percent of the median, the median, and 120 percent of the median. We will use this wage distribution information from four scenarios: (1) "wage expectation conditional on being of age 30 and having achieved secondary education as highest education", i.e., leaving the University of Applied Sciences now,[5] (2) "expectation conditional on age 40 and having achieved secondary education as highest education", (3) "expectation conditional on age 30 and having achieved tertiary education" and (4) "expectation conditional on age 40 and having achieved tertiary education".

The information we got from the students does not fully identify the underlying wage distri-

[5] The vast majority of students has completed a commercial apprenticeship (*Berufslehre*) on upper secondary level (ISCED 3B) before entering the University of Applied Sciences.

butions. By assuming a specific distribution function, we can retrieve its defining parameters and calculate any moment of the distribution, such as variance and skewness. Enforcing a certain distribution comes, however, at a cost: Every distribution has its own features which limit the way students' expectations can be represented. Fitting a log-normal distribution, as Dominitz and Manski (1996) and Wolter (2000) do, imposes a heavy restriction on the set of possible student expectations. The two-parameter log-normal distribution is, among other features, always positively skewed.

It seems highly unlikely that all students should have such a distribution function in mind for all the scenarios. This can easily be shown by looking at the share of distributions that are positively skewed: only 62 percent of the 1008 distributions elicited are positively skewed. Assuming a log-normal distribution is thus not precisely correct for roughly one third of the individual distributions.

There is another reason why the assumption of log-normality seems too restrictive. Tsiang (1972) shows that risk averse individuals appreciate positive skew. Hence, we would like to test whether students do indeed expect compensation for wage risk and a subtraction for skewness at the same time. The log-normal distribution does, however, not allow us to separate mean, variance and skewness: it is fully described by the parameters mean and variance, so skewness cannot vary independently from these parameters. Assuming log-normal distributions, we implicitly assume that students do not distinguish between variance and skewness when making expectations. Thus, we could not test whether positive skew is associated with a lower expected mean wage, conditional on expected wage risk.

We will therefore not only fit log-normal distributions[6] to the wage distribution information elicited from students and use the interquartile range (iqr) of these distributions as a measure of risk. Instead, we will also specify alternative, non-parametric measures of variance and risk. The three pieces of information we ask students about their expected wage distribution - the median and the probabilities associated with one value below and one value above the median

[6] We also fitted Beta distributions instead of log-normal distributions. The literature (see McDonald 1984) typically finds that the Beta distribution performs better than the log-normal distribution in fitting wage distributions since it entails two shape parameters (instead of one in case of the log-normal). Applying a root mean squared error criterion, our Beta distributions perform worse, however, than the log-normal in three of the four scenarios. The reason is that some students gave answers that indicate distributions that are strongly skewed to the right. The log-transformation is well suited to deal with these cases, whereas the Beta distribution parameters take on extreme values without providing a good fit. The log-normal seems to be the best parametric assumption available for our purpose, despite its shortcomings noted in the text.

- can be used to define simple variance and skewness measures. The three points divide the respective probability density function into four parts. We denote the probability masses lying in the four intervals by A, B, C and D respectively:[7]

$$\begin{aligned} A &= P(0 \leq w < 0.8 * m) \\ B &= P(0.8 * m \leq w < m) \\ C &= P(m \leq w < 1.2m) \\ D &= P(1.2 * m \leq w < \infty) \end{aligned}$$

By definition of the median (m) we know that $A + B = C + D = 0.5$. Then a natural variance measure is defined by looking at the share of total probability that has been assigned to the two outer parts of the distribution: $v = A + D$. This provides us with a non-parametric *variance coefficient* (not to be confounded with the "coefficient of variation") that lies between 0 and 1. In the same vein, a *skewness coefficient* can be defined by looking at the asymmetry in the probabilities assigned to the two outer parts of the distribution: $s = 2(D - A)$. This coefficient lies between -1 and 1; a positive sign indicates positive skewness and vice versa, while 0 indicates a symmetric distribution.[8,9] Figure 1 illustrates how students' answers are reflected by the various variance and skewness measures. It shows three hypothetical cases, where every student expects a median wage of 5000 CHF and was thus asked for the probability to earn less than 4000 or more than 6000 CHF. Student 1 expects lower wage risk than student 2, but both allocate probabilities symmetrically. By contrast, student 3 has the same wage risk as student 2, but her expected wage distribution is positively skewed.

Our non-parametric measures seem intuitively appealing and less restrictive than assuming log-normal distributions and deriving variance measures from them. It remains, however, to be decided how the interval width around the median, i.e. the boundary between the probability areas A and B and between C and D should be defined. They can be defined to be

[7] Because of rounding to digestible values, the boundaries are not in all cases located at exactly $\pm 20\%$ from the median, so minor adjustments of the respective probabilities were necessary, see Appendix C.2.

[8] Of course, the limited information available about the density functions does not allow us to identify higher or lower variance and skewness unambiguously. Implicitly, we are still making distributional assumptions.

[9] Note that these definitions do not imply correlations between the variance coefficient and skewness from the presence of the same term in both (A and D), as a change in A must always imply a change in B, C and/or D, without preset pattern. Indeed, the correlation turns out to be virtually zero (see section 2.3).

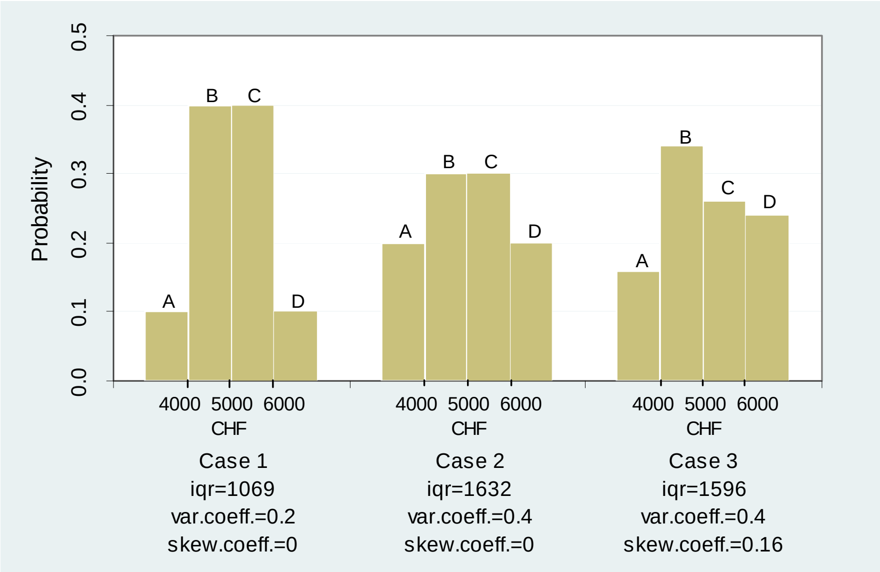

Figure 1: Three hypothetical wage distributions of students with expected median 5000 CHF and fitted variance and skewness measures

proportional to the median (i.e., ±20% around the median), or as an interval with fixed width for all scenarios and persons. This depends on the type of risk aversion of students' utility functions.[10] Risk averse students are indifferent between receiving the expected value of the (expected) wage distribution for sure and getting a draw from that wage distribution plus a certain risk premium. If students exhibit constant absolute risk aversion, the risk premium they expect for wage risk depends only on the variance, not on the expected value of the wage distribution. In this case, a fixed interval width independent of the median seems the best choice as basis for the calculation of a variance coefficient.

By contrast, if students exhibit constant relative risk aversion, they expect a risk premium which is constant for risk that is proportional to their wealth, independent of their wealth level. In other words, the risk premium is constant for the wage variance divided by the expected value of the wage distribution. Defining the variance coefficient based on a variable interval width growing and shrinking proportionally to the median seems more adequate in this case.

We will use both specifications and compare the results. The *relative interval width* specification uses the interval defined in the survey, i.e. ±20% around the median, to determine the

[10] For the following short discussion on risk aversion, we assume that students' wealth is largely determined by their life time income from work.

probability masses A to D. The *fixed interval width* specification, however, implies that the interval endpoints, and thus the probabilities A to D, have to be adapted for each observation, in order to ensure that the same interval is used for everybody in every scenario. We used the mean interval width in the pooled sample which defines the new interval endpoints at 1530.9 CHF above and below the respective median. The probability mass lying between the original (relative) and the new endpoints was moved accordingly.[11] This requires assuming a distribution function. We have used the log-normal distribution again. After these adjustments, the variance and skewness coefficients as defined above were computed for the fixed interval width specification.

Although the proposed variance and skewness coefficients do in principle not require the strong assumption of log-normality, the fixed interval specification had to make use of this assumption to a certain extent. Nonetheless, our alternative variance measures are quite different from using the interquartile ranges of the fitted log-normal distributions. In addition, we are able to calculate skewness measures and to assess skewness independently from variance.

2.3 Descriptives for risk and skewness measures

One may worry that the probabilistic information we draw from students is not "spontaneous" information but is stamped and moulded by the nature of the questionnaire: the feedback in the computer programme reduces misunderstanding by simply not allowing violation of basic rules on probability. First, we may repeat that Dominitz and Manski (1996) focussed precisely on this problem and that they concluded that the method yields meaningful information. Second, we have kept track of interventions in cases where individuals specified a probability above or below the median surpassing 50 percent. An individual has to specify 20 probabilities for a total of 10 scenarios. He can thus potentially make 20 errors, although this would indicate utter incomprehension, as intervention would never improve understanding. 65 percent of the respondents never made any error, 20 percent made one error, 11 percent made two errors and the remaining 4 percent made more than two errors (the worst was a single person with 7 errors). That does not seem to suggest that the answers depend critically on the error

[11] For instance, if the interval relative to the median was wider than the fixed interval, probability mass had to be moved from B to A and from C to D. The adjustments for the fixed interval width specification were as follows: in 2016 adjustments (252 students * 4 scenarios * 2 wage values around median), 41.4% of all adjustments included changes of more than 5 percentage points. 7.7% entailed changes of more than 10 percentage points.

correction: the data are basically obtained from a sample of respondents who did not need more than a single reminder. Finally, we have used this information for a robustness check and we found no indication of a serious flaw in our results (see 3.2).

Figures 2 to 4 present the distributions of expected median wages as well as variance and skewness coefficients for the four scenarios. The distributions are based on the 252 cases in the sample each, those presented in figure 3 and 4 are defined on a fixed interval width.

Figure 2: Distribution of median of students' expected wage distributions

As Figure 2 shows, the distribution of medians for secondary education at age 30 is fairly concentrated. The dispersion of expected medians across individuals increases with level of education and with age. For tertiary education, at age 40, the distribution has a remarkably long upper tail.

Figure 3 shows that distributions of the variance coefficients are quite different for the different scenarios. Variances are clearly higher for the scenarios at age 40 and for the tertiary education scenarios. While the distribution of the variance coefficient is strongly skewed for the scenario age 30/secondary education, with hardly any values above 0.5, the distribution for scenario 40/tertiary education appears almost symmetric around 0.5. It is obvious that students assign

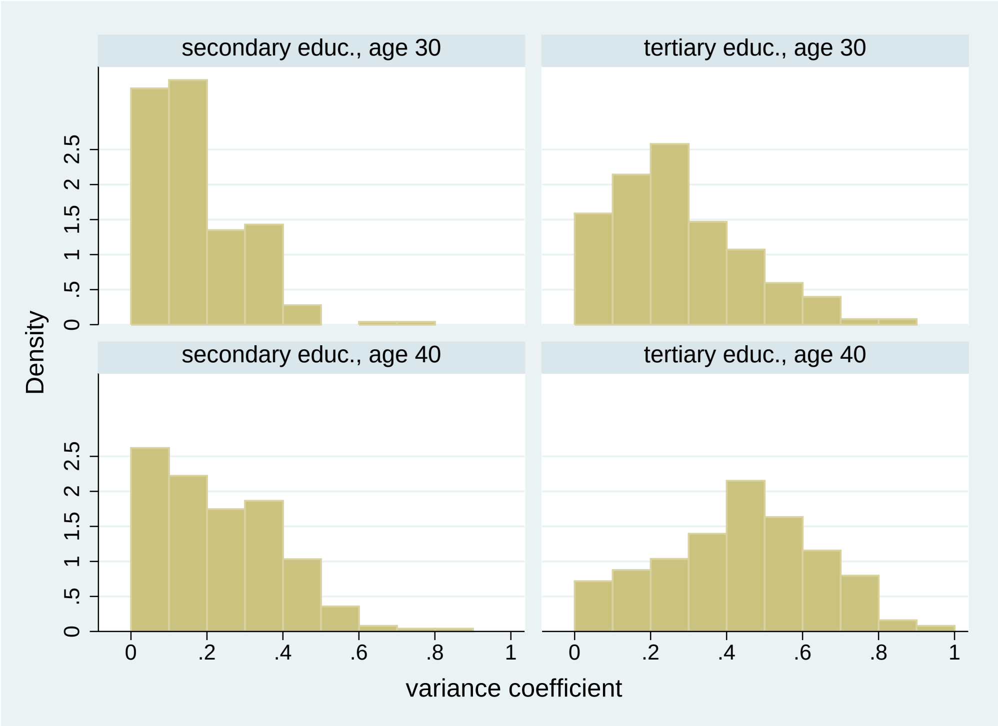

Figure 3: Distribution of variance coefficients of students' expected wage distributions

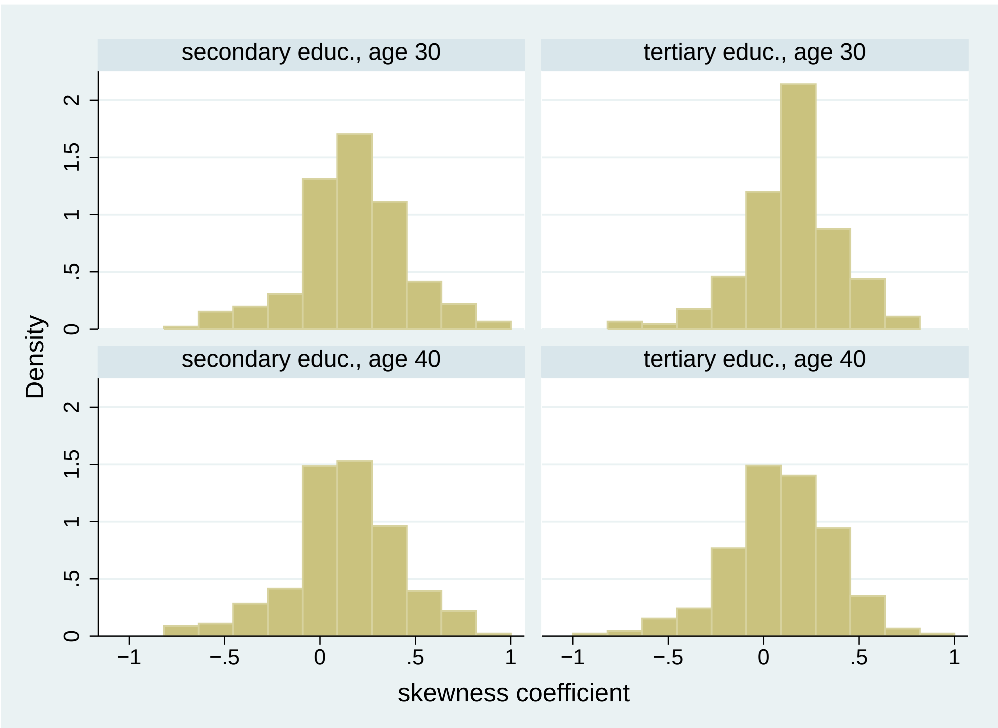

Figure 4: Distribution of skewness coefficients of students' expected wage distributions

much more wage risk to the latter scenario than to the former.[12]

For skewness, the picture is less clear. The distributions of the skewness coefficients seem a bit broader for the scenarios with age 40, but actually, the dispersion of skewness coefficients is remarkably stable across scenarios; locations also vary modestly, with the lowest mean for 40/tertiary. To compare the scenarios in more detail one needs to consult the descriptive statistics in the Appendix. The most interesting message of figure 4 is that expected skewness varies considerably between individuals. An important share of the expected wage distributions is negatively skewed. Furthermore, skewness is only loosely correlated with variance: Calculating the correlation of variance and skewness conditional on scenario dummies, skewness has a marginally significant positive correlation with variance. The partial correlation is, however, very small (regression coefficient 0.04). Most of the variance in skewness is not driven by the variance (risk) of the underlying wage distributions. These findings confirm that log-normality is not a fully satisfactory approximation for all individual wage distributions.

We conclude that individuals' anticipated wage distributions exhibit much variation, are not always symmetric (nor symmetric in logs) and exhibit variation in skewness independent from variation in the individuals' anticipated variance.

3 Results: students expect compensation for wage risk

3.1 Core results

If workers are risk averse, they should be compensated for wage risk and a higher risk should lead to a higher mean wage. Thus, one way to assess the importance of wage risk is to estimate Mincer earnings equations including a measure of wage risk r and skewness s:

$$ln(w_i) \quad = \quad X_i\beta_x + \beta_r r_i + \beta_s s_i + \epsilon_i \qquad (1)$$

In the literature (see Hartog and Vijverberg 2007 for an application), risk r_i has been measured

[12] Not surprisingly, the differences between the scenarios are smaller if one considers the variance coefficient based on a variable wage interval around the median (see descriptives in the Appendix). The risk in scenario 30/secondary education appears higher then, whereas it appears lower in scenario 40/tertiary education. The ranking of scenarios remains the same, however, with the highest risk attached to scenario 40/tertiary education.

as the variance around the mean wage in the particular group to which individual i belongs, i.e. education or occupation. The argument is that individuals build wage expectations for alternative educations and occupations by just looking at the wage distributions they observe on the labour market for the particular groups. The variance around the mean, within schooling-education groups, is a measure of the individual's ignorance, of the unpredictability of wages and hence, of risk. Typically, the regression also contains a measure for the skewness s_i of the wage distribution within the occupation/education group: just as expected wages for some education should increase with the variance because individuals dislike risk, the expected wages may be lowered for positive asymmetry in the distribution. Risk averse individuals appreciate a long upper tail of the distribution as it gives them favourable odds of large gains relative to large losses (Tsiang 1972), and they are willing to pay for it by accepting lower wages, thus exhibiting skewness affection. Different authors (see King 1974; McGoldrick 1995; McGoldrick and Robst 1996; Hartog and Vijverberg 2007 for the US, and Hartog et al. 2003 for Europe) have chosen the risk augmented Mincer approach and have found that mean income in an occupation or education is positively related to the variance and negatively to the skewness.[13]

Typically, the equation is estimated in two steps, with variance and skewness defined on the residuals (within occupations/educations) from an ordinary Mincer equation in the first stage and then added to a re-estimation in the second stage. The main criticism applying to this approach is that ex post wage realizations are not a valid proxy for the ex ante wage risk r_i (and skewness s_i) faced by the agents (Cunha et al. 2005; Cunha and Heckman 2007). Only part of the variance that can be found in actual wage data is due to risk, another part is due to worker heterogeneity. Heterogeneity means that individuals have superior knowledge compared to the researcher who looks at ex post data. Individuals would then not just look at the average wages they observe on the labour market for different groups. If individuals have private information about their own ability and other productivity-related variables, they form more informed expectations about where they will end up in the wage distribution. Researchers who do not have this information would then overestimate individuals' wage risk when looking at the variance of ex post realizations of wages. Cunha et al. (2005) and

[13] The mentioned papers did not analyze the case of Switzerland. We have replicated their work with data of the Swiss Labour Force Survey and find the same qualitative results as these authors: variance has a negative sign, skewness a positive sign in a risk augmented Mincer earnings regression. Detailed results are available from the authors.

Cunha and Heckman (2007) promote this argument and present an econometric solution for the problem. They develop and apply a method for decomposing cross section variability of earnings into components that are forecastable at the time students decide to go to college (heterogeneity) and components that are unforecastable (risk). Instead of reconstructing the information set from observed behaviour, as Cunha and Heckman did, we will exploit the direct observations on expectations, i.e. students' own forecasts. In particular, we can see if individuals' expectations reflect the risk compensation by using the expected variance and skewness measures in Mincer earnings equations. We will start with pooled results, i.e., the data for 4 scenarios for each of 252 individuals has been combined, giving 1008 cases.

Table 1 shows OLS regression results for the dependent variable log median wage. Column I shows a wage regression without risk and skewness measures. According to the scenario dummies, students expect to earn 24 percent more at age 40 than at 30 if they would go working immediately.[14] Completing tertiary education, they think to earn 39 percent more at age 30 than without tertiary education. At 40, they expect another 45 percent on top of that when finishing tertiary education. Year dummies reflect the boom in 2000/2001. Men expect somewhat higher wages than women. Higher wages for men and steeper age-wage profiles for higher education are stylised facts that students are clearly aware of.

Different risk and skewness measures are the variables of interest in the models II to VI. In column II, this is the interquartile range derived from the fitted log-normal wage distributions. We find a significant positive effect on median wage, which mirrors the findings with data from actual, ex post wage realizations. The mean expected effect of risk on wages is substantial: An increase in the interquartile range by 1,000 CHF is associated with an expected increase of earnings by 4 percent. This result is of comparable magnitude to results in the US, where McGoldrick (1995, p.221) found that "a $1,000 increase in the standard deviation of unsystematic earnings [the risk measure; note from the authors] will increase men's earnings by 2.5 percent and women's earnings by 3.1 percent".[15]

The inclusion of a risk measure increases the goodness of fit of the estimation and has an effect on other coefficients. The scenario dummy coefficients are reduced in some specifications, meaning that also the expected return on tertiary education becomes lower. In Hartog

[14] The results on scenario dummies in the text refer to the exact percentage changes (i.e., $e^{\hat{\beta}} - 1$) instead of the values approximated by the dummy coefficients in the table.

[15] The exchange rate was 1.182 CHF/USD in 1995 (source: Swiss National Bank).

Dep. var.: ln expected median wage	I	II	III	IV	V	VI
interquartile range (divided by 1000)		0.040**				
		(0.014)				
Variance coeff. (fixed interv. width)			0.441**	0.446**	0.440**	
			(0.046)	(0.047)	(0.047)	
skewness coeff. (fixed interv. width)				−0.073**	−0.073**	
				(0.024)	(0.023)	
Variance coeff. (relative interv. width)						0.134**
						(0.049)
skewness coeff. (relative interv. width)						−0.061*
						(0.024)
Scenario age 40/secondary education	0.219**	0.201**	0.187**	0.184**	0.184**	0.213**
	(0.006)	(0.009)	(0.007)	(0.008)	(0.008)	(0.007)
Scenario age 30/tertiary education	0.327**	0.297**	0.276**	0.273**	0.274**	0.320**
	(0.009)	(0.014)	(0.011)	(0.011)	(0.011)	(0.010)
Scenario age 40/tertiary education	0.612**	0.542**	0.495**	0.487**	0.488**	0.596**
	(0.011)	(0.024)	(0.015)	(0.017)	(0.016)	(0.012)
year 1999	0.049*	0.037	0.031	0.025	0.032	0.044+
	(0.024)	(0.024)	(0.022)	(0.022)	(0.022)	(0.024)
year 2000	0.103**	0.088**	0.067**	0.066**	0.076**	0.098**
	(0.028)	(0.026)	(0.024)	(0.024)	(0.023)	(0.027)
year 2001	0.136**	0.125**	0.109**	0.112**	0.124**	0.144**
	(0.026)	(0.024)	(0.024)	(0.024)	(0.024)	(0.026)
Age	−0.004	−0.004	−0.005	−0.005+	−0.004	−0.004
	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Male	0.063**	0.043*	0.027	0.027	0.027	0.050*
	(0.020)	(0.019)	(0.017)	(0.017)	(0.016)	(0.019)
part time study	0.035	0.051	0.076*	0.063*	0.063+	0.043
	(0.036)	(0.033)	(0.032)	(0.032)	(0.033)	(0.037)
father's education high					0.020	0.003
					(0.036)	(0.038)
father's education low					0.005	0.004
					(0.016)	(0.018)
mother's education high					0.013	0.025
					(0.027)	(0.031)
mother's education low					−0.005	0.013
					(0.034)	(0.034)
upper class					0.106+	0.124+
					(0.056)	(0.065)
upper middle class					0.037*	0.049**
					(0.017)	(0.019)
lower class					−0.048	−0.056
					(0.039)	(0.045)
Second. school grade French					0.014	0.003
					(0.024)	(0.027)
Second. school grade German					0.051*	0.054*
					(0.024)	(0.027)
Second. school grade Math					0.014	0.011
					(0.012)	(0.014)
Intercept	8.531**	8.515**	8.512**	8.531**	8.096**	8.149**
	(0.092)	(0.082)	(0.075)	(0.074)	(0.169)	(0.183)
F-Test	400.22	443.73	476.77	422.46	223.26	182.36
Adj. R-squared	0.654	0.693	0.718	0.723	0.735	0.678
N	1008	1008	1008	1008	1008	1008

Significance levels: + p<0.10, * p<0.05, ** p<0.01;
Standard errors are corrected for clustering of students due to pooling. Reference group: scenario age 30/secondary education, year 1998, female, father's education medium, mother's education medium, middle class

Table 1: Earnings regressions using individual wage expectations (OLS)

et al. (2003, table 1), the education variable remained unaffected by the risk and skewness variables. Our differing result could have important implications concerning the interpretation of expected, ex ante rates of return to education, as part of the ex ante return may have to be re-interpreted as risk compensation. However, this effect depends on using fixed or relative interval width for calculating the variance coefficient.

Column III presents the results for the variance coefficient described in section 3.2 which is used in place of the interquartile range in column II. Again, we find a significant and positive effect on the median wage. This effect hardly changes when the skewness coefficient is added (column IV).

Adding controls for individual background (column V) does not influence the coefficients for either variance or skewness. Thus there are no spurious correlations or biases if omitted. An increase in the variance coefficient from 0 (which means that all probability mass has been assigned to the interval plus/minus 1'530 CHF around the median) to 1 (the full probability mass is assigned to the lower and upper end of the distribution[16]) is associated with a more than 40 percent higher median wage. Although this calculation is based on the maximum possible difference in variance, the order of magnitude shows that the effect is substantial even for smaller variance differences.

Column VI shows the results with the variance (and skewness) coefficient defined on an interval width proportional to the median. The result is qualitatively the same, although the coefficients' size as well as the goodness of fit are reduced.

The skewness coefficient shows a negative sign and is significant in all models, though its effect is clearly weaker than that of risk. A higher skewness is associated with a lower median. As discussed in the introduction to this section, this can be explained by students' risk aversion which implies skewness affection.

Using expectation (i.e. ex ante) data, we can thus fully replicate the results of the literature on risk augmented Mincer earnings equations which uses actual ex post wage data: expected risk variables show a positive effect, expected skewness variables a negative effect on expected median wage. In fact, we even get similar values for the elasticities. Multiplying the regression coefficient with the mean values of risk and skewness (0.285 and 0.137, respectively) we find a risk elasticity of 0.125 and a skewness elasticity of -0.010 (for fixed interval width), values that

[16] This case is theoretical and means an infinite variance.

are in the middle of the interval of values found in the empirical literature. The elasticities for the case of variable interval width are lower (at 0.040 and -0.009, respectively), but still within the range found for market wages (see Hartog 2007 for a summary of elasticities estimated on market data).

3.2 Robustness checks

Different objections might be raised against our interpretation of the results in table 1. We will discuss the following three possible shortcomings in turn: a) the wage expectation data might be unreliable, b) pooling across scenarios might hide heterogeneous results across the scenarios, and c) there might exist unobserved heterogeneity across students that biases the results.

a) Unreliable wage expectation data?

As pointed out in section 2.1, the expectation data is of high quality due to the computer assisted interactive survey. Our software did, however, not only point out inconsistencies and errors to respondents, it also traced these errors - which again, is an advantage over paper and pencil survey data. We can therefore include variables for the number and type of errors respondents have committed. These refer to misunderstandings of the concept of probability and the median, i.e. stating probabilities higher than 100 percent or stating probabilities higher than 50 percent for the parts of the distribution above or below the median. Including this information on errors in the regressions of table 1 does not influence the results; neither does the exclusion of the (small) share of people who committed several errors. Given the high data quality, the plausible descriptive results of the survey and the stability of the results using different specifications in table 1, we are confident that our results are not an artefact caused by unreliable data.

b) Does pooling obscure heterogeneity in scenarios?

Pooling the observations for four different scenarios per person might hide heterogeneous results for regressions run separately for every scenario. Four separate regressions all show the same results for our variables of interest (positive sign for the variance coefficient, negative for the skewness coefficient), however (not shown).

c) Unobserved heterogeneity across students?

Although we control for different individual characteristics in the regressions of table 1, there might still exist student fixed effects, i.e. unobserved student characteristics that are correlated with expected median as well as with expected risk and skewness.

Dep. var.: ln median wage		
variance coeff. (fixed interv. width)	0.356**	
	(0.030)	
skewness coeff. (fixed interv. width)	−0.054**	
	(0.015)	
variance coeff. (relative interv. width)		0.120**
		(0.035)
skewness coeff. (relative interv. width)		−0.045**
		(0.016)
scenario age 40/secondary education	0.191**	0.214**
	(0.009)	(0.009)
scenario age 30/tertiary education	0.284**	0.321**
	(0.009)	(0.009)
scenario age 40/tertiary education	0.512**	0.598**
	(0.012)	(0.010)
Adj. R-squared	0.836	0.808
N	1008	1008

Significance levels: + p<0.10, * p<0.05, ** p<0.01;
Reference group: scenario age 30/secondary education

Table 2: Fixed effects estimation of risk augmented earnings equations

Therefore, we estimated a fixed effects model where the students' means over the four scenarios have been subtracted from each variable. All variables that are fixed for a student drop out of the estimation. Table 2 therefore only includes scenario dummies in addition to the variance and skewness coefficients.

The results are in line with the results of the comparable models IV and VI in table 1, although the coefficients for variance and skewness are slightly reduced.

4 Conclusions

We have investigated students' expectations on the benefits of education, by using data from direct questioning rather than deduced from imposed econometric modelling. This provides us with individual ex ante wage risk measures that are not confounded with heterogeneity as is the case with using group variances of actual market wages as a risk measure. Students anticipate wage distributions with substantial variance and skewness: they are far removed from anticipating a single wage for a particular schooling choice as assumed in the standard human capital model. Although expected dispersion differs widely between students, on average students anticipate the variance to increase with age and to be higher with tertiary than with secondary education. These expectations are in line with actual labour market data. We further find that students are aware of risk compensation, with implicit elasticities remarkably close to those actually observed in the labour market. Earnings uncertainty is thus an important part of the uncertainty inherent in individuals' educational decisions, with students expecting a trade-off between risk and return as it is known from the stock market.

References

Betts, Julian R. (1996). What Do Students Know About Wages? Evidence from a Survey of Undergraduates. *Journal of Human Resources 31(1)*, 27–56.

Brunello, Giorgio, Claudio Lucifora, and Rudolf Winter-Ebmer (2004). The wage expectations of European college students. *Journal of Human Resources 39(4)*, 1116–1142.

Cunha, Flavio and James Heckman (2007). Identifying and Estimating the Distributions of Ex post and Ex ante Returns to Schooling. *Labour Economics 14(6)*, 870–893.

Cunha, Flavio, James Heckman, and Salvador Navarro (2005). Separating uncertainty from heterogeneity in life cycle earnings. *Oxford Economic Papers 57*, 191–261.

Dominitz, Jeff and Charles Manski (1996). Eliciting Student Expectations of the Return to Schooling. *Journal of Human Resources 31*, 1–26.

Friedman, Milton (1953). The Methodology of Positive Economics. In MiltonFriedman (Ed.), *Essays in Positive Economics*. Chicago: University of Chicago Press.

Hartog, Joop (2007). A Risk Augmented Mincer earnings equation? Taking stock. *Working Paper Universiteit van Amsterdam*.

Hartog, Joop, Erik Plug, Luis Diaz-Serrano, and Jose Vieira (2003). Risk compensation in wages - a replication. *Empirical Economics 28*, 639–647.

Hartog, Joop and Wim Vijverberg (2007). On Compensation for Risk Aversion and Skewness Affection in Wages. *Labour Economics 14(6)*, 938–956.

King, Allen G. (1974). Occupational choice, risk aversion and wealth. *Industrial and Labor Relations Review 27*, 586–596.

Kodde, David A. (1986). Uncertainty and the Demand for Education. *Review of Economics and Statistics 68(3)*, 460–467.

Levhari, David and Yoram Weiss (1974). The Effect of Risk on the Investment in Human Capital. *American Economic Review 64(6)*, 950–963.

Manski, Charles (2004). Measuring Expectations. *Econometrica 72(5)*, 1329–1376.

McDonald, James B. (1984). Some generalized functions for the size distribution of income. *Econometrica 52(3)*, 647–664.

McGoldrick, Kimmarie (1995). Do women recieve compensating wages for earnings risk? *Southern Economic Journal 62*, 210–222.

McGoldrick, Kimmarie and John Robst (1996). The effect of worker mobility on compensating wages for risk. *Applied Economics 28*, 221–232.

Nicholson, Sean (2002). Physician Specialty Choice under Uncertainty. *Journal of Labor Economics 20(4)*, 816–847.

Tsiang, Sho-Chieh (1972). The Rationale of the Mean-Standard Deviation Analysis, Skewness Preference, and the Demand for Money. *American Economic Review 62(3)*, 354–371.

Webbink, Dinand and Joop Hartog (2004). Can students predict their starting salaries? Yes! *Economics of Education Review 23(2)*, 103–113.

Weiss, Yoram (1972). The Risk Element in Occupational and Educational Choices. *Journal of Political Economy 80(6)*, 1203–1213.

Wolter, Stefan C. (2000). Wage expectations: A comparison of Swiss and US Students. *Kyklos 53*, 51–69.

Wolter, Stefan C. and André Zbinden (2001). Labour Market Expectations of Swiss University Students. *International Journal of Manpower 23*, 458–470.

Zafar, Basit (2007). College major choice and the gender gap. *Department of economics, Northwestern University*.

A Descriptives (N=252)

variable	**mean**	**standard deviation**
year 1998	0.274	
year 1999	0.198	
year 2000	0.187	
year 2001	0.341	
age	23.6	2.49
male	0.706	
part time study	0.099	
father's education high	0.071	
father's education middle	0.830	
father's education low	0.099	
mother's education high	0.032	
mother's education middle	0.357	
mother's education low	0.611	
upper class	0.028	
upper middle class	0.369	
middle class	0.540	
lower class	0.063	
Second. school grade French[a]	4.88	0.435
Second. school grade German[a]	4.97	0.368
Second. school grade Math[a]	4.80	0.639

[a]Maximum grade is 6, minimal passing grade is 4.

Table 3: Individual fixed variables (no variance between scenarios)

variable \ scenario	age 30/secondary education		age 40/secondary education		age 30/tertiary education		age 40/tertiary education	
	mean	std. dev.	mean	std. dev.	mean	std. dev.	mean	std. dev.
expected median wage (CHF)	5294	842	6619	1228	7346	1220	9852	2209
interquartile range (log-normal distrib.; CHF)	1085	657	1532	905	1826	1131	2810	2316
variance coefficient (fixed interv. width)	0.172	0.124	0.245	0.166	0.288	0.175	0.437	0.2
variance coefficient (relative interv. width)	0.263	0.159	0.289	0.163	0.301	0.166	0.342	0.171
skewness coefficient (fixed interv. width)	0.176	0.283	0.137	0.291	0.152	0.254	0.083	0.279
skewness coefficient (relative interv. width)	0.189	0.289	0.141	0.296	0.154	0.254	0.099	0.282

Table 4: Variables varying with scenarios

B Phrasing of questions on wage expectations in the questionnaire

The following questions on students' own wage expectatios were introduced by a section explaining the meaning of probabilities and the median, and by specifying details about the wages asked (per month, full-time equivalent, no inflation).

Scenario secondary education:
Imagine you stop studying now and do not start another education. Think about the kind of occupations, industries, hierarchy levels etc. in which you will be working under these conditions. What is the median amount of money that you think you will earn by the time you are 30 (40) years old? What do you think is the probability that you will earn more than X / less than Y? At age 30: ... / At age 40: ...

(Original German version: Stellen Sie sich vor, Sie brechen Ihr jetziges Studium ab und absolvieren keine zusätzliche Ausbildung. Beantworten Sie die folgenden Fragen, indem Sie sich Berufe, Branchen, Hierarchiestufen etc. vorstellen, in denen Sie unter dieser Voraussetzung arbeiten würden. Wie hoch schätzen Sie Ihren Medianlohn im Alter von 30 bzw. 40 Jahren ein? Wie schätzen sie die Wahrscheinlichkeit ein, dass Ihr Lohn höher wäre als X / niedriger wäre als Y? Mit 30 Jahren: ... / Mit 40 Jahren: ...)

Scenario tertiary education:
Imagine you have successfully finished your education at the University of Applied Sciences before age 30. Think about the kind of occupations, industries, hierarchy levels etc. in which you will be working under these conditions. What is the median amount of money that you think you will earn by the time you are 30 (40) years old? What do you think is the probability that you will earn more than X / less than Y? At age 30: ... / At age 40: ...

(Original German version: Stellen Sie sich vor, sie haben die Ausbildung an der Fachhochschule vor dem 30.Lebensjahr absolviert. Beantworten Sie die folgenden Fragen, indem Sie sich Berufe, Branchen, Hierarchiestufen etc. vorstellen, in denen Sie unter diesen Umständen arbeiten würden. Wie hoch schätzen Sie Ihren Medianlohn im Alter von 30 bzw. 40 Jahren ein? Wie schätzen sie die Wahrscheinlichkeit ein, dass Ihr Lohn höher wäre als X / niedriger wäre als Y? Mit 30 Jahren: ... / Mit 40 Jahren: ...)

C Calculation details

C.1 Fitting log-normal distributions

The log-normal distribution is completely determined by two parameters, typically expressed as the mean μ and standard deviation σ of the underlying normal distribution. The log of the median m of the log-normal distribution equals μ by definition of the log-normal distribution: $\ln(m) = \mu$. The interquartile range iqr of the log-normal distribution can be calculated by estimating σ. Since the median and two additional points of the wage distribution are known, σ can be estimated. Starting from the z-transformation

$$\frac{x-\mu}{\sigma} = z(p),$$

rearranging and substituting the known log wages for x and μ gives:

$$ln(w) - ln(m) = \sigma * \Phi^{-1}(p)$$

We have two observations for w and p each, such that σ can be estimated as the coefficient in an OLS with N=2 using the equation above (adding an error term at the right hand side). With σ at hand, the interquartile range of the fitted log-normal distribution can be computed:

$$iqr = m * (e^{z_{.75} * \sigma} - e^{z_{.25} * \sigma})$$

where $z_{.75} = \Phi^{-1}(0.75) = -z_{.25}$.

C.2 Probability adjustments due to rounding of the interval endpoints

In order to compare the measures between persons, the definition of the intervals containing probabilities A to D has to be the same across persons. Because the values 0.8 times median and 1.2 times median had been rounded off to the nearest 500, the interval defined by these values does not have a width of exactly 40 percent of the median. Moreover, the interval becomes asymmetric depending on the median. For instance, a median of 6,100 results in a lower value of 5,000 (instead of 6,100*0.8=4,880) and in an upper value of 7,500 (6,100*1.2=7,320). Both values in this example have been rounded up.

We defined new endpoints of the interval that are exactly 0.8 times median and 1.2 times median. Then, the probability mass lying between the original (rounded) and the new endpoints had to be moved. In the example above, probability mass had to be moved from A to B and from C to D in order to find the probabilities associated with the values that equal exactly 80 percent and 120 percent of the median. This requires assuming a distribution function. We have used the log-normal distribution again. The actual size of probability adjustments (for the specification using interval width relative to the individual median) was minor: in 2016 adjustments (252 students * 4 scenarios * 2 wage values around median), only 16 cases occurred where more than 5 percentage points of probability had to be moved. After these adjustments, the variance and skewness coefficients presented in 2.2 have been computed.