A Risk Augmented Mincer Earnings equation? Taking stock

joop hartog

Universiteit van Amsterdam

in cooperation with

Simona Bajdechi Raita, Peter Berkhout, Jose Cabral Vieira, Luis Diaz Serrano, Bas Jacobs, Helena Skyt Nielsen, Hans van Ophem, Erik Plug, Juerg Schweri, Wim Vijverberg, Dinand Webbink

JEL code: J31; D8

Abstract

We survey the literature on the Risk Augmented Mincer equation that seeks to estimate the compensation for uncertainty in the future wage to be earned after completing an education. There is wide empirical support for the predicted positive effect of wage variance and the negative effect of wage skew. We discuss robustness of the findings across specifications, potential bias from unobserved heterogeneity and selectivity and consider the core issue of students' information on benefits from education.

First draft: 16 October 2006;this draft 20 November 2006
Comments received in the joint UvA-VU Labour seminar and at CREB (University of Barcelona) are gratefully acknowledged.

1. Basic hypothesis

There can be no doubt that the decision to engage in an education is a decision under uncertainty. There are at least three dimensions in which information is incomplete. The potential student will generally not fully grasp the requirements of the school curriculum and of the occupations available after graduation, will usually not even be fully sure of her own abilities and preferences and will not know what the exact returns to the investment will be. With uncertainties so prominent, one is inevitably led to expect that the returns to education will be shaped not only by compensation for postponing earnings, but also by compensation for risk. In fact however, the basic Mincer earnings equation is derived under conditions of certainty on future earnings in the alternatives and the estimations similarly ignore the impact of uncertainty. Research on returns to education has focussed for decades on getting an unbiased estimate of the causal effect of schooling on earnings. Main emphasis is on the proper econometric modelling to deal with omitted ability bias, measurement errors in education and the endogeneity of the schooling decision. Admittedly, these are serious and stubborn problems, as illustrated by the fact that the dust has not settled down and that there is as yet no consensus on true, causal, rates of return to education and their variation in response to individual and institutional conditions. Perhaps all we can be confident about is an interval of the return for an additional year of education between a few percent to an upper limit of some 20 percent.

With uncertainty so prominent, one might have thought that there is a large literature to deal with it. In fact, analyses investigating the consequences of uncertain completion of an education, incomplete information on abilities and even preferences and on the poor predictability of the career development that comes with an education are very scarce. This paper can only make a small contribution, by treating just one aspect of the pervasively present uncertainty. We focus on the notion that an individual contemplating an education does not anticipate a given post-school wage rate but rather an entire wage *distribution*, without knowing exactly where in that distribution she will end up. It turns out that a basic idea on an econometric approach to this problem has been around for a long time[1]. The literature has started with King (1974), who used aggregate data by occupation to estimate the effect of variance and skew of wages within an occupation on the mean occupational wage. The variance of earnings is taken as a measure of risk that individuals want compensation for. Positive skewness of the earnings distribution points to the opportunity of attaining real high earnings and this is something people want to pay for with reduced expected earnings. With micro data one may proceed in two steps. First, estimate a standard Mincer earnings equation, group the residuals by education-occupation classification and take the within-group distribution of residuals as indicating the uncertainty associated with choosing the particular education-occupation combination. In the second stage, add the variance of these education-occupation residuals to the earnings function and estimate risk compensation as the regression coefficient of earnings on the variances. The two-stage approach was first applied by McGoldrick (1995). Following King (1974), she also included the skew of residuals by occupation in the regression equation.

[1] After I suggested this approach to Luis Diaz Serrano for his dissertation, he dug up the early references. A referee on one of the papers discussed here noted that the same idea had been repeatedly formulated by students.

Compensation for earnings risk has now been estimated in close to twenty studies. Generally, the studies report an earnings premium for risk and a rebate for skew, as theory predicts. In this paper we review these studies and assess the evidence so far. In section 2 we present a formal model underlying the estimations, in section 3 the empirical specifications. In Section 4 we present estimation results. Section 5 addresses the problems of bias from heterogeneity and presents estimates that control for these problems. The thorny issue of causality is taken up in section 6: for several reasons one may think the relationship is spurious, caused by other factors than market reaction to the risk that potential students face. Section 7 concludes by assessing the results obtained so far and outlining further work. .

2. A simple formal model.

We can formally derive the required compensation for risk from imposing equal expected lifetime utility for all educations. For the sake of exposition, we assume that there is one option that has fixed earnings in every year that an individual works; these earnings are known to the individual. We ignore experience effects for individuals in all options. We also ignore compensation for postponing earnings when going to school, as this is taken care of in the usual Mincer mark-up (and simply added on in our empirical specification).

We start with the simplest possible model, modifying just one assumption in the classic Mincer framework. A potential student can choose between going to work right way and obtain a known and fixed income for the rest of his working life, or go to school and upon graduation get a draw from an earnings distribution that will be constant until retirement[2]. As we assume identical students, market equilibrium will only emerge if both options generate equal expected lifetime utility: expected wages in the uncertain alternative should compensate for risk. In the riskless alternative, annual earnings are given as Y_f, generating utility $U(Y_f)$, where $U(\ )$ is a concave utility function with $U' > 0$, $U'' < 0$ and $U''' > 0$ (the latter condition is necessary for declining absolute risk aversion[3], see Tsiang, 1974 or Hartog and Vijverberg, 2002). In the risky option, income is a single draw for the rest of working life, written as Y_r+e. Equal expected lifetime utility requires

$$\int_0^T U(Y_f)e^{-rt}dt = E\int_0^T U(Y_r+e)e^{-rt}dt \tag{1}$$

where T is the length of working life and r the time discount rate. We can write the left-hand side as

$$\int_0^T U(Y_f)e^{-rt}dt = \frac{1}{r}\left(1-e^{-rT}\right)U(Y_f) \tag{2}$$

For the stochastic term on the right-hand side we apply a third-order Taylor expansion around the expected value Y_r, one order up from Pratt's original contribution (Pratt, 1964), to

$$\int_0^T U(Y_r+e)e^{-rt}dt = \frac{1}{r}\left(1-e^{-rT}\right)\left[U(Y_r)+\frac{1}{2}U''(Y_r)s_p^2+\frac{1}{6}U'''(Y_r)k_p^3\right] \tag{3}$$

where σ_p^2 is the second moment (risk) and κ_p^2 is the third moment (skewness) of e around the expected value zero. Equating (2) and (3) and rewriting a little, after applying a first-order Taylor expansion around Y_r for (2), we get

---

[2] In Hartog and Vijverberg (2006) we estimate parameters of the utility function in a model with new shocks in every year of experience. In Diaz Serrano et al (2006) we estimate compensation using panel data and we extend static model (4) with annual "transitory" shocks.

[3] As pointed out in Tsiang (1974), increasing absolute risk aversion is unrealistic. Therefore, we should expect that a risk-averse individual displays preference for skewness in addition to aversion to dispersion of the probability distribution of returns, so that, we should expect that $U_y''' > 0$.

$$\frac{Y_r - Y_f}{Y_r} = -\frac{1}{2}\frac{s_p^2}{Y_r^2}\frac{U''}{U'}Y_r - \frac{1}{6}\frac{k_p^3}{Y_r^3}\frac{U'''}{U''}Y_r\frac{U''}{U'}Y_r = \frac{1}{2}\frac{s_p^2}{Y_r^2}V_r - \frac{1}{6}\frac{k_p^3}{Y_r^3}V_s V_r \tag{4}$$

where V_r is Arrow-Pratt's relative risk aversion and V_s is the similar definition for relative skewness affection (we call it affection, because individuals like skewness; for evidence from other choices see Garrett and Sobel, 1999 and Golec and Tamarkin,(1998); in lifecycle consumption-savings modeling, skew affection is called prudence, see Gollier, 2001). With V_r and V_s positive by definition, we note from (4) that individuals only enter an education if the utility loss from uncertainty is matched by a positive premium for the risk (variance), while they allow an earnings drop for skewness. Thus, equation (4) predicts that wages respond positively to risk and negatively to skew when individuals decide on education based on their knowledge of the second and third moment of the wage distribution associated with an education[4].

By necessity, the model is a simplification. It is an extension, with stochastic rather than deterministic post-school earnings, of the framework specified by Jacob Mincer, still the basis for all routine estimates of returns to education. Mincer (1974, 9-11) did not even bother to spell out the strong underlying assumptions on market structure, individual abilities and information, except for a brief reference in the introduction. More elaborate estimates of the causal effect of schooling on earnings are often based on Card (1999). Card's model endogenises schooling, allows for unobserved heterogeneity in individuals'

[4] Cochrane (2001) points out that in a lifetime welfare maximising framework variance as such is not relevant: it's the covariance of an asset with consumption that is essential. Levhari and Weiss also stress the key importance of the covariance between marginal utility of consumption and returns to education. We still denote the residual variance of wages as risk, as we abstract from optimal consumption profiles over the life cycle.

cost and benefits but does not allow for any imperfect information by individuals: they know all their cost and prospective returns with certainty. Card's problem is an ignorant researcher looking at an omniscient individual. His model serves his purpose, but it is not the single best model to serve all purposes. Clearly, we make very strong assumptions. But so did Mincer and so did Card. In fact, our model is quite similar to Levhari and Weiss (1974), the seminal paper on human capital investment under uncertainty. They also use a two period model with uncertain second period returns to investment made in the first period. As shown in Appendix 1, our basic equation can also be derived from their model.

A key assumption of our approach is that individuals cannot insure the risk of their investment. To us, this is obvious. We simply do not observe individuals commencing a college education and buying insurance or an optimal investment portfolio that completely eliminates the risk of their venture. Davis and Willen (2000) compute optimal portfolios for some occupations and find completely unrealistic values. For example, a 40-year old truck driver (in 1982) should hold a portfolio of $550,000, including a short position in one portfolio of $141,000. Our view is shared by e.g. Blanchard and Fisher (1989:283) and by Shaw (1996:626) who states: "The methods of reducing riskiness that are available in financial markets, namely, diversification, exchange, and insurance, are not options for reducing the riskiness of returns to human capital investments". Palacios-Huerta (2003) studies the relationship between human capital risk and financial investment in a lifecycle consumption framework and reports supporting evidence. He finds that at the aggregate level, the mean-variance frontier does not improve if returns from financial assets are added to returns from human capital, whereas in the converse case (adding human capital to financial assets) the frontier does improve. For separate demographic groups, the results vary by level of education. Shaw (1996) reports a similar result, based on her own analysis and reference to earlier work: the covariance between human and financial wealth is zero, leaving no scope to reduce human capital risk by adequate financial investment.

Note that in the simple formulation of the model given above, the only uncertainty is the post-school wage rate. It's hard to see how this risk can be insured. Of course there is also substantial uncertainty about the further development of earnings during the individual's career. We will discuss this career uncertainty later in the paper. One might argue that an individual does not have to accept uncertain earnings over the lifecycle as an inescapable event, as it is always possible to apply a minimum degree of consumption smoothing.

3. Empirical specifications

Suppose, risk aversion V_r and skew affection V_s would be constant, as would hold for the CRRA utility function $U(Y) = \frac{1}{1-r} Y^{1-r}$, implying $V_r = r$ and $F_r = r + 1$; then equation (4) would hold exactly, and we could use it to estimate the parameter of the utility function.

If we add continuous discounting with discount rate d we can add the Mincerian compensation for postponing earnings (see Hartog and Vijverberg, 2002), to obtain

$$E(\ln Y_s) = \ln Y_o + \frac{d}{1-r} s + \frac{1}{2} r \frac{m_{2s}}{m_s^2} - \frac{1}{6} r (r+1) \frac{m_{3s}}{m_s^3} \qquad (5)$$

which is a simple equation in schooling years, variance term and skew, with m and m the empirical counterparts of the terms in (4). Hence with observations on relative variance and relative skew we could estimate the *Risk Augmented Mincer equation* (RAM). If we don't assume CRRA, the parameters will not be constant.[5] Equation (5) is still a good starting point for empirical work regressing earnings on the parameters of the distribution. Several specifications have been used to obtain measures of risk and skew, and below we will present a survey of estimation results.

The first analysis, by King (1974), used variance and skew of earnings in occupational cells and estimated equation (5) at the aggregate level. Feinberg (1981) used a short panel to estimate each individual's coefficient of variation over 6 years. McGoldrick (1995) introduced the two-step procedure. With cross-section data, she first estimated an earnings function

$$\ln Y_{ij} = X_i b + \sum_j a_j d_j + e_{ij} \tag{6}$$

where the subscripts i and j denote individuals and the education cell the individual belongs to respectively. The d_j are dummy variables for education cells (fixed effects). The variables included in X are years of education, age and age squared and, depending on specification, dummies for gender and ethnicity. Generally, no other explanatory variables in X are included, as the common variables that may be available (such as industry, firm and job characteristics) are all unknown to the individual when deciding on

[5] Moore (1995) derives a similar equation for risk compensation (variance) under uncertainty on the present value of lifetime earnings and CRRA preferences.

education. The education fixed-effects a_j are included in order to control for the effect of omitted variables that may bias the measures of risk and skew within an education cell. Estimated residuals are used to compute measures of R and K

$$R_j^{(1)} = \frac{1}{N_j}\sum_i \left(e_{ij} - \bar{e}_j\right)^2 \qquad\qquad K_j^{(1)} = \frac{1}{N_j}\sum_i \left(e_{ij} - \bar{e}_j\right)^3 \tag{7}$$

where e_{ij} is the exponential of the estimated residuals e_{ij} in equation (6). In (7), R and K are simply estimated as the second and third moment of the distribution of *exp(e$_j$)*. In the second step the estimated values for R and K are included in the wage equation

$$\ln Y_{ij} = X_i b + g_R R_j + g_K K_j + e_{ij} \tag{8}$$

Dummies for education cells cannot be included in (8) since R and K are already fixed in a given education cell.

Hartog and Vijverberg (2002) introduced a specification that is closer to the model in section 2. First, estimate

$$LnW_{ji} = X_i b + e_{ji} \tag{9}$$

where i indicates the individual and j indicates the occupation-schooling group that the individual belongs to. Years educated is one of the variables in the matrix *X*. Define s_j^2 as the variance of the disturbance e_{ji} in occupation/education cell j. Use the estimated parameter vector $\hat{b}$ and the estimated variance $\hat{s}_j^2$ to predict the wage rate for each individual through:

$$\hat{W}_{ji} = \exp\left(X_i \hat{\beta} + \hat{\sigma}_j^2 / 2 \right) \tag{10}$$

Finally, calculate wage deviations $W_{ji} - \hat{W}_{ji}$ and from these the relative variance R_j and relative skewness K_j , defined as

$$R_j = \frac{1}{I_j} \sum_{i=1}^{I_j} \left(\frac{W_{ji} - \hat{W}_{ji}}{\hat{W}_{ji}} \right)^2 \tag{11}$$

$$K_j = \frac{1}{I_j} \sum_{i=1}^{I_j} \left(\frac{W_{ji} - \hat{W}_{ji}}{\hat{W}_{ji}} \right)^3 \tag{12}$$

In (10), the variance term is added to the mean to reflect that the disturbances of the earnings distributions are approximately lognormal, as is commonly assumed. Were the distribution indeed lognormal, equation (10) would hold exactly.[6] R and K are the sample estimates of relative variance and relative skew. In practice, (7) and (11)-(12) are equivalent: in the Danish data (Diaz-Serrano, Hartog and Nielsen, 2003) the measures in (11) and (12) correlate better than 0.99 in each of 17 years with those in (7).

4. Estimates of the Risk Augmented Mincer Equation

In Table 1, we present estimates of the Risk Augmented Mincer Equation. *Grosso modo* one could say these are first generation estimates, where risk and skew are measured within occupations. Whenever relevant, the regressions include a parabolic age (experience) profile. Sometimes there are additional controls, in particular in the second

---

[6] Hartog and Vijverberg (2002: 2004) test for log normality and mostly reject it. Still, adding the variance reduces the bias in the estimate of the mean.

stage (the first stage should not control for effects that the individual cannot anticipate when choosing an education, such as e.g. firm size). The estimates strongly support the basic hypothesis of risk compensation. The coefficient for risk is positive in all but one case and significant, usually at high levels in all studies except the study with union interaction. The elasticity is mostly in the interval 0.1 – 0.2. The coefficient of skew is negative, except in 4 out of the 19 cases, usually but not always at high levels of statistical significance. The elasticity is small, mostly below 0.10. The study with union interaction, by Moore (1995), focuses on benefits that unions bring their members. Compensation for wage risk is insignificant in both unionized and non-unionised jobs. Interestingly, in jobs covered by union contract workers have less wage variability and more hours variability: unions appear to prefer fixed-wage contracts with variable hours. They also bring higher compensation for wage risk than in the non-union sector, but the difference is statistically not significant.

The early studies and the replications suffer from flaws that have to be addressed before any firm conclusions can be drawn. First, earnings risk is measured at the level of occupations, and thus will be sensitive to selective mobility. Individuals who are not successful in an occupation may try their luck elsewhere[7]. With overrepresentation of workers with good draws, observed earnings overestimate the earnings of all those who tried this occupation. With truncation of the earnings distribution at the low end, observed variance of earnings is an underestimate of risk, while skew is overestimated. Thus the risk coefficient will be overestimated. The effect on the coefficient of skew

[7] Johnson (1977) and McGoldrick and Robst (1996) report a higher compensation in occupations with less mobility to other occupations: lock-in effects are compensated.

cannot be predicted as both expected earnings and skew are overestimated. Better estimates can be obtained with observations grouped by education, as one cannot escape bad draws after the education has been completed. If risk induced mobility between occupations is important, one must predict that education based estimates give lower

Table 1 The Risk Augmented Mincer equation: occupations

Author	country	year	n	R, K	? (R)	?(K)	controls
King (1974) Table 2, row 3							
	USA	1960 Census	37 occupations (professions, 4 years college)	Standard deviation. Third moment	1.22[a] (8.71)	-4.44 E-3[a] (4.77)	Ability (Project Talent)
Johnson (1977)							
	USA	1970 Census	55-107 occupations	Standard deviation	0.11[b)]	-	Within age-education groups
Feinberg (1981) Table 1, column (1)							
	USA	1971-1976 PSID panel	1419 individuals	Individual coefficient of variation over 6 years	0.01 (2.53)	-	IQ, education, occupation higher compensations for the more risk averse
McGoldrick and Robst (1996), p. 230							
	USA	PSID 1979-1984	528 women 937 men	Standard deviation residual earnings (time, time sq)	M: 0.17 (11.57 F: 0.50 (6.87)		Mobility, 7 occupations
Ma (2005)							
Table 8.6	China	1991 Urban Household Survey	41 education-occupation	Exp (?)	M:0.05 (1.90) F: -0.23 (8.14)	0.05 (7.92) -0.006 (0.96)	
Table 8.5		2000	51 educations-occupations	Exp (?)	M:0.18 (3.70) F: 0.65 (10.75)	0.02 (0.80) -0.09 (3.26)	
Hartog et al (2003) Table 1,2							
	Netherlands	OSA 1999	40 occupations	Exp (?)	M:0.18 (10.02) F:0.15 (3.96)	-0.05 (8.18) -0.06 (3.32)	6 industry dummies
	West Germany	SOEP 1992	41 occupations	Exp (?)	M: 0.11 (4.02) F: 0.25 (6.06)	-0.02 (3.07) -0.05 (4.42)	6 industry dummies
	East Germany	SOEP 1992	28 occupations	Exp (?)	M: 0.12 (5.03) F: 0.08 (2.64)	-0.01 (0.91) -0.02 (1.57)	6 industry dummies
	Portugal	Quadros Pesoal 1992	70 occupations	Exp (?)	M: 0.37 (33.35) F: 0.76 (30.65)	-0.17 (24.08) -0.03 (2.32)	6 occupation dummies, 7 industry dummies, tenure, firm size, ownership and age

Table 1 continued

Author	country	year	n	R, K	? (R)	? (K)	controls
Hartog et al (2003) Table 1,2							
	Spain	Estructurra Salarial 1995	83 occupations	Exp (?)	M: 0.29 (58.11) F: 0.10 (11.65)	-0.02 (8.45) -0.02 (35.83)	Bargaining regime, public/private, 6 industry dummies, 6 occupation dummies, city size
Hartog and Vijverberg (2006) Table 3							
	USA	NBER-CPS 1995-1999	129 education-occupation	Relative variance, skew	M: 0.16 (15) F: 0.05 (4)	-0.06 (26) -0.17 (14)	Elasticities, t-values averaged
			104 education-occupation	Interquantile ranges	M: 0.44 (20) F: 0.08 (3)	-0.26 (12) -0.45 (15)	
Diaz-Serrano (2000) Table 7.3							
	Spain	EPF 1990	83 occupation	Exp (?)	M: 0.02 (7.4) F: 0.04 (4.2)	-0.02 (2.8) -0.10 (4.8)	Family status, region, industry, skill level
Moore (1995) Table 4							
	USA	PSID 1978-1987	856 individuals	Individuals coefficient of variation	Union –0.08 (0.39) Non-union 0.09 (?)		Risk interacted with union dummy

[a] Regression coefficients: units and means not reported

[b] Relative effect of one standard deviation on mean earnings, averaged over 18 age-education categories. All coefficients significant at conventional levels; t-values refer to estimated coefficient, not to elasticity

Table 2 The Risk Augmented Mincer equation: education

Author	country	year	n	R, K	? (R)	? (K)	controls
Diaz-Serrano, Hartog, Nielsen (2004) Table 2							
	Denmark	1984-2000	75 educations	Exp (?)	0.03[a)]	-0.005[a)]	Men, aged 30-40 clustered
Berkhout, Hartog and Webbink (2006) Table 1, all							
	Netherlands	LSO 1997	66 educations	Exp (?)	0.2 (3.69)	-0.1 (2.57	clustered
-	-	Elsevier/SEO 1996-2001	100 educations	Exp (?)	0.08 (4.53	-0.04 (3.25)	Starting salaries, tertiary education clustered
Diaz-Serrano and Hartog (2006)							
Table 3, Model 1	Spain	1995	53 educations	Exp (?)	0.222 (16.50	-0.034 (2.42)	clustered
Model 2				relative error	0.202 (16.96)	-0.008 (10.00)	
Model 3				Interquartile ranges	0.035 (11.67)	-0.016 (2.50)	
Berkhout and Hartog (2006)							
	Netherlands	Elsevier/SEO 1996-2006	111 educations	Exp (?)	M: 0.05 (2.19) F: 0.04 (2.43)	-0.13 (5.50) -0.10 (6.43)	Regression on mean earnings by education, personal characteristics clustered; starting salaries
Hartog and Vijverberg (2007)							
	USA	NBER-CPS 1995-1999	DOT-SEO cells	Cell means	M: 1.07 (9.34) F: 0.86 (5.41)		Region, ethnicity, working conditions

[a)] Elasticity averaged over 17 annual estimates (all significant); mean R and K from IZA DP 963 estimated risk compensation than occupation based estimates, while the effect on estimated skew affection cannot be predicted.

The studies in Table 1 also fail to recognize a complication in the estimated standard errors. As noted by Moulton (1986), with R and K measured at group level, errors may be correlated within these groups and estimated standard errors should be corrected for this clustering within cells[8].

Table 2 presents second generation estimates, with risk and skew measured by education and with standard errors corrected for clustering, or with mean earnings regressed on mean risk and skew. The basic conclusion is unaffected: a positive effect of risk and a negative effect of skew, both statistically significant. We cannot compare the magnitudes of the coefficients (elasticities) between Table 1 and Table 2, as they are based on different datasets and different controls, so we cannot test the prediction that education based measures lead to lower elasticities than occupation based measures. As we noted above, there is no difference between measuring risk and skew with the exponential specification or as relative measures (equation (7) versus (11)-(12)). We have also tested another specification, with R and K based on interquartile ranges. As R and K will be sensitive to outliers, we have calculated the percentile distribution of the residuals in the first stage, and defined R as the difference between the 75the and the 25th percentile and K as $(P_{75} - P_{50})/(P_{50} - P_{25})$. As can be seen in both in Table 1 and in Table 2, we still obtain significant effects with the proper signs.

[8] The Moulton problem is not relevant for Johnson (1977) and King (1974), who estimated on aggregate data by occupation and for Feinberg (1981), who estimated on panel data.

On top of the Moulton problem, we have to realise that R and K are measured from a first stage regression and will be subject to sampling errors. Murphy and Topel (1986) have presented a method for consistent estimation of the variances of the second stage parameter estimates. In Diaz Serrano and Hartog (2006), the corrections for clustering and for generated regressors are combined, by replacing the conventional OLS covariance matrix estimator in Moulton's adjustment by the covariance estimator proposed by Murphy and Topel. The Moulton correction for clustering is important, but the additional correction for generated regressors has negligible effect on estimated standard errors.

With R and K fixed for a given occupation or education, second stage regressions cannot include a fixed effect. Thus, it may be that R and K do not measure any dimension of risk but simply some fixed effect of occupation of risk that is not picked up by the controls. In Diaz Serrano, Hartog and Nielsen (2003) we have regressed education fixed effects from the first stage earnings functions on R and K (both permanent and transitory, see below). The effect of R and K on these fixed effects was not statistically significant, thus indicating that R and K are not just representing education fixed effects through the back door. In Diaz Serrano and Hartog (2006) we applied the same test to our estimates for Spain, with the same result. We can be pretty confident that R and K are not just unidentified education effects

5. Unobserved heterogeneities

An immediate objection to using the observed distribution of earnings residuals is the possible confounding of risk and heterogeneity. The (residual) distribution of earnings will be affected by many variables that for an individual may not pose any risk at all. Probably the most prominent of these variables is individual ability. But individuals may also differ in risk attitudes and they may even face individually different risk. To address these issues, in Jacobs, Hartog and Vijverberg (2005) we set up a simple model with a safe fixed wage job (with wage W_s) and a risky job with stochastic wage, at expectation $E(W_r)$. We did not refer specifically to educations as the problem is identical in a situation where individuals can choose a job or an occupation without first spending time in school (the required Mincer compensation for postponing earnings can simply be added to earnings).

If individuals are identical in all aspects (risk aversion r, risk s^2 and "ability" expressed as expected productivity in the risky job), in equilibrium the expected risky wage will just compensate for risk and allocation will be arbitrary. Risk is properly estimated as observed wage variance in the risky job, the coefficient of risk compensation, estimated as the observed wage gap divided by wage variance, without bias, from the equilibrium condition $E(W_r) = W_s + r s^2/2$. This is a stripped down version of the model in section 2. Note that r is estimated from the wage gap divided by half the variance.

To assess the effects of heterogeneity, we distinguish three cases. First, we assume that individuals only differ in risk attitudes, and have identical abilities and risk. This poses no special problem. The wage gap between risky and safe jobs is now determined by the risk attitude of the marginal individual, and this is just what is estimated in the OLS regression. Second, if individuals only differ in the magnitude of risk and are otherwise identical, residual earnings variance underestimates true risk, as only low risk individuals enter. As a consequence, risk compensation is overestimated. If individuals differ in both risk and risk attitude, the case is rather unusual. The threshold for entering the risky job is now defined as a critical value of the product of risk and risk attitude: the threshold for entry is a contour of combinations of the two variables that separate entry from non-entry. There is no single critical value of any of these variables and by consequence they cannot be estimated. The coefficient that is estimated in the Risk Augmented Mincer equation is now the maximum value of risk aversion for the individual that has the mean value of risk among those individuals who actually chose the risky job. The third case is where individuals only differ in ability, reflected in their expected earnings in the risky job; assuming individuals know their ability and the researcher does not, we have a garden variety of selectivity. The observed wage gap between the two jobs is now an overestimate of the wage premium, as the mean value of the risky wage also includes the effect of ability for those who chose the risky job. Observed wage variance in the risky job is an overestimate of risk, as it also includes the variance of ability between individuals. With risk aversion estimated as the ratio between wage gap and observed variance, we cannot predict the sign of the bias, as both numerator and denominator are biased upward. To assess this case, we revert to simulation. We find that the ambiguity in the sign of the bias cannot be resolved by restriction to reasonable or plausible parameter values: depending on parameter values it may just as well be positive as negative. Thus of course implies that one cannot dismiss the risk compensation estimates as unreliable because they always produce an overestimated coefficient.

Below, we will discuss extensively whether ability is known to the individual and may generate an unobserved heterogeneity for the researcher, or whether individuals are also poorly informed. Here, we just note that observations on ability are available in several datasets, and that we are in a position to assess the effect of omitted ability bias in our estimates. Interestingly, the very first tests of the risk compensation hypothesis already allowed for ability differences. King (1974) controlled for ability as measured in the Project Talent. Significant effects of risk and skew were found while controlling for five ability measures: mathematics, English, reading comprehension, abstract reasoning and arithmetic reasoning. The basic conclusion was also upheld after splitting the sample in two classes of competing groups, technical and non-technical (a distinction confirmed by discriminant analysis on the basis of the five abilities). Estimation separately within the two groups, with their own measures of risk and skew, confirmed the basic results. Feinberg (1981), testing on a panel with individual intertemporal variance as his measure of risk, included IQ (a sentence completion test, on which no further information is given).

In Berkhout, Hartog and Webbink (2006) we analyse starting salaries for graduates from tertiary education in The Netherlands. Instead of using R and K for all individuals with the same education, we calculate these measures for ability quartiles as measured by the individuals' average grades for their secondary school final exam. This is a relevant stratification, as it is based on information for which we are sure that the individuals themselves also have it. The model underlying the risk compensation hypothesis assumes that students perceive their risk and then decide on education. If desired, they could use the same information as we do, by stratifying on secondary school grades. As the results in Table 3 indicate, after controlling for ability in this way the basic results are upheld. In conformity with our analysis we see that the bias from omitted ability can go either way. For vocational education, coefficients increase a little if we control for ability, but for university education they are cut in half. The latter result suggests that the overestimation of risk is larger than the overestimation of the mean.

Table 3. Effect of controlling for ability on estimated coefficients (Elsevier/SEO)

	R	T	K	t	N
All tertiary					
Total	1.69	4.53	-0.022	3.25	31 893
By quartile	1.20	4.22	-0.013	2.79	31 893
Vocational					
Total	0.90	1.61	-0.004	0.51	14 955
By quartile	0.99	3.28	-0.007	1.74	14 955
University					
Total	2.18	6.67	-0.026	2.69	16 938
By quartile	1.24	3.14	-0.013	1.83	16 938

By quartile: R and K measured for individual's quartile for secondary school exam grades. Source: Berkhout, Hartog and Webbink (2006); clustered standard errors

Panel data are interesting for at least two reasons. By considering individual wage variation over time, unobserved heterogeneity is eliminated. Some of the first generation studies used panel data, and measured risk as the individual variance over time. Feinberg (1981) and McGoldrick and Robst (1996) each used 6 waves of the PSID and found statistically significant effects of individually measured risk, although the magnitude of the effect in the former study was very small. Moore (1995) measures risk individually from ten waves of the PSID and finds no significant risk compensation within either the union or the non-union sector, but instead finds that unions provide wage insurance (lower variance) at the price of higher hours variability.

Second, with panel data we can separate earnings variation between individuals and for given individuals over time. With a large and long Danish panel (1984-2000) we separate the earnings residual in an individual (random) effect and an annual transitory shock. We take the variance of the individual effects of individuals within the same education as the permanent risk of that education and the variance of the transitory component within an education as transitory risk. We insert the values in the wage equation for the pooled panel and we report the results in Table 4. The compensation for permanent and for transitory risk are both highly significant, though quite modest in magnitude. As to skewness, compensation for the permanent component is highly significant, with the right sign. For the transitory component, however, the prediction of a negative sign is squarely rejected, with a significant positive effect. The results for the permanent component are at variance with the hypothesis that individuals know their own fixed effect and need no compensation for it.

Comparable results are obtained with the NLSY panel for the United States, with the same risk decomposition (permanent risk from individual random effects, transitory risk from the annual transitory shocks), but with compensation for risk only tested on the wage equation for the finale year. Again we find that risk compensated has the predicted sign, both permanent and transitory, although not significantly so for permanent. For skewness, the predictions find less support: only transitory skewness is significant with the right sign. Interestingly, if we calculate risk and skewness by IQ deciles, the results are virtually the same: omitted ability (as measured by IQ) does not do much harm.

Table 4: Wage compensation for transitory and permanent shocks

Denmark [a] 1984-2000		
	Risk	Skewness
Permanent	0.3322	-0.0481
	(17.90)	(17.84)
	0.0390	*-0.0104*
Transitory	1.5727	4.1831
	(6.00)	(8.74)
	0.0472	*0.0128*
USA [b] 1979-2000		
Permanent	0.001	0.076
	(0.99)	(1.32)
Transitory	0.309	-0.976
	(12.05)	(11.94)
By ability decile		
Permanent	0.010	0.002
	(1.92)	(1.15)
Transitory	0.297	-0.106
	(2.55)	(3.99)

Notes:
a. Estimates include years of education, age, age squared, and dummies for industry and occupation. Each cell contains coefficient, t-value in parentheses and elasticity in italics. Compensation estimated on pooled panel data. Source: Diaz Serrano, Hartog and Nielsen (2004)

b. Controls included for marital status, region, some occupation dummies, IQ test score (AFQT); the data are from NLSY 1979 (1979-2000). Compensation estimated on wages in 2000. Source: Raita, 2005, Tables 5.6, 5.7.

6. Supporting evidence

To consider the credibility of the results, we may consider some circumstantial evidence. First of all, we may note that there is nothing unusual about risk compensation. Compensation in the stock market is of course commonplace, but risk compensation in the labour market is also well established, both for instability of employment and for health and morbidity hazards (Rosen, 1986). In fact, employer behaviour exhibits the mirror effects of the risk compensation we analyse here. Fresh graduates pose a risk for employers to the extent that they cannot accurately assess their potential productivity. But as graduates have no alternative for putting their education at work, we expect employers to shift the risk of unknown abilities to employees. Using variance and skewness of students' grades in different disciplines as indicators of employer uncertainty we find that starting wages are lower in fields with high variance and higher in fields with high skewness (Berkhout and Hartog, 2006).

More specifically, we may point out that there is evidence that individuals indeed care about the financial risk when choosing a labour market position. King (1974) and Saks and Shore (2003) note that with risk aversion declining in wealth, one should expect that students from wealthier backgrounds choose more risky occupations. This is precisely what they find, with American data. One might extrapolate this to choice of education, but not much evidence is available. Kodde (1985), using Dutch data on individual students' expectations, finds no clear effects of the gap between highest and lowest earnings expected with a university on probability to attend. Lazear (2005) reports that individuals are more likely to start their own business firm if the first industry in which they were employed after leaving university has high wage variance, implying they have relatively low risk aversion.

Hartog and Vijverberg (2006) is the only paper that attempts to estimate structural parameters in the wage compensation equation. The estimated coefficient of relative risk aversion is 0.64 for men and 0.46 for women, which is in the low end of the interval of estimates from various sources, such as consumption –savings models, TV games and direct surveys. The coefficient of relative skewness affection is estimated at 1.03 for men and 2.18 for women, higher than the only other value we found in the literature (0.3, in a consumption-savings model). The interesting observation is that the implied preference parameters are very much in line with similar values found in quite different applications, and that attitudes towards financial risk in the labour market are not fundamentally different from attitudes revealed in consumption-saving decisions.

One may have doubts that R and K properly measure the uncertainty they are meant to measure. In that respect it is reassuring that if we use dispersion measures derived from the percentile distribution of residual earnings, the results are not essentially different. Supporting evidence is also provided by Pereira and Martins (2002). They measure risk as the difference between returns to education in the highest and the lowest decile in quantile regression. Regressing the Mincer rate of return on risk across 16 countries, they find a significant positive effect with implied elasticity of 0.2. Even though there are only 16 observations, the evidence is neatly in line with the other results we report here.

Shaw (1996) argues that on-the-job training is a risky investment and that less risk averse individuals will invest more. She exploits the fact that the share of financial wealth held in risky assets is proportional to the degree of risk aversion (and hence, that this share can be used as an index of risk aversion) but also uses a direct survey question on risk aversion[9]. She finds indeed that less risk averse individuals have higher wage growth. If the share of financial wealth invested in risky assets increases from the sample mean of 0.10 to 0.4, her model implies that the share of human capital devoted to investment in creating new human capital would also increase fourfold. The 3-year growth rate of the wage would then increase from 6.6% to 7.8 %, i.e. an increase by 18%. Her results can also be used to calculate that compensation for earnings risk as emphasized in this paper has an elasticity of 0.5: moving from the class of above-average risk takers to the class of average risk takers, residual log wage variance falls 50% while the average wage falls 30%, moving to the class of no-risk takers, residual log wage variance falls 70%, while the average wage falls 40%. Shaw's results provide strong evidence in favour of our approach. Her results are not based on *assumed* risk taking behaviour, they hold for individuals who differ in stated risk attitudes. Individuals who indicate that they take more financial risk to obtain larger gains have higher average wage and higher residual wage variance. The same relationship holds in growth rates: less risk averse individuals

[9] Her model is very similar to the model later developed by Hartog and Vijverberg (2007) to analyse school curriculum choice as an optimal risk-return portfolio.

have higher wage growth rate and higher variance of residual wage growth[10]. Replication of Shaw's model on Spanish data (in progress) confirms these results.

Diaz-Serrano (2005) and Dohmen et al. (2007) document that individuals sort themselves into labour market positions according to their risk attitudes. Diaz-Serrano measures earnings uncertainty as residual variance in Italian panel data after removing individual fixed effects and finds that earnings uncertainty correlates significantly with the stated reservation price for a lottery ticket: more risk averse individuals have less earnings uncertainty. Dohmen et al use a measure of individual risk attitude obtained from a survey question and find clear support for this measure in a validation exercise with experimental data. They measure earnings risk as residual earnings variance by occupation and report significant correlation between risk and individual risk attitude: more risk averse individuals sort into jobs with less earnings variance.

7. Should we correct for selectivity?

The hypothesis of risk compensation in wages is based on the assumption that supply decisions by potential students affect market wages. Students are supposed to react on the financial uncertainty associated with an education. The estimations we presented use observed residual variances as an indicator of risk. It's by now a natural reflex of labour economists to assume a selectivity problem here, on the hypothesis that individuals are well informed on their abilities and their income prospects (see e.g. Chen, 2006; Chen

[10] The effect of the increase in the risky asset share is given in Shaw's Table 2. Changes in mean wages by risk class are given by the regression results at the bottom of page 641. Variances by risk class are given in footnote 22.

and Khan, 2005). Selectivity corrections are based on the assumption that students, when they take their decisions, have information on future returns that the researcher can only observe in later realisations for the same set of potential students. However, there are good reasons to be critical on the assumption that potential students are so well informed. Direct indicators of students' expectations are particularly negative. Before entering this discussion, it is good to point out that we are *not* interested in the true risk associated with an education. What matters in our case is the *belief* of potential students on the risks they are facing. As usual in economics, there are two approaches: one may impose a model structure and estimate the apparent beliefs, as revealed by observed behaviour, or one may apply a direct approach and elicit individuals' beliefs by interviewing them. The difference relates to different methodological stances, but should preferably be solved by empirical evidence.

Dominitz and Manski (1996) have shown that students are highly uncertain on their benefits from an education. They interviewed high school and college student on their earnings expectations under different scenario's. Their study is rather exploratory in methodology, with only 71 high school students and 39 college students, but it is very carefully executed and gives unusual insights. Students have widely divergent anticipations. For example, among male high school students asked for expected median earnings at age 30 in the scenario where they will have completed a bachelor degree, the 10th decile expects 25 000 dollars whereas the 90^{th} decile expects 56 000 dollars. Students were also interviewed on the dispersion in their own future distribution of earnings. For example, the same groups, under the same scenario, expect interquartile ranges of 28 000 and 58 000 dollars (interquartile ranges are derived from questions on the probabilities to surpass specified thresholds)[11]. Perceived returns differ widely. At age 30, the male high school students at the 10th percentile expect an earnings advantage from college of 10 000 dollars, at the 90th percentile they expect a gain of 30 000 dollars. Students were also asked to state the actual dispersion of earnings by education. Generally, they overestimate the interquartile range. Interestingly, predictons of their own median expected salary correlate positively with their perception of the actual median: "Respondents who believe current median earnings to be high (low) tend also to expect their own earnings to be high (low)" (o.c., p 25). Dominitz and Manski do not relate their results to family background, personal ability or school achievements, so we cannot rule out that the large variation expected interquartile ranges (which we might label risk) is related to superior information. But the positive correlation between expected personal median and perceived actual median is compatible with students using actual distributions as a basis for their own expectations, with perception errors carrying over into variation in their expectations. Perception errors may be unrelated to their own qualities. Results by Brunello, Lucifora and Winter-Ebmer (2004) can be interpreted as support for this view. They collected information on expected earnings from university students in ten European countries (business and economics). The expected wage premium over high school graduates at labour market entry was unrelated to any variable except age: not to parental background, not to channel of information about future earnings (university publication, career center, special reports, press, personal communication), not to reason for choosing their selected university, not to self-assessed

[11] The answers on medians and ranges have no fixed patterns. Across groups and scenarios, the answers and their ratios differ widely.

relative ability. The expected premium of university over high school graduation after ten years of experience is only significantly lower for older students, women and students with longer expected time to complete the degree (and again not to self-assessed ability). These results only refer to university students, not to those who choose not to go to university, but they do not point to systematic patterns in expected benefits, except for the effect of expected length of study, which possibly reflects an update of initial expectations during the study.

Wolter and Weber (2000) apply the Dominitz and Manski approach to Swiss students. They also find large dispersion in expected medians (a coefficient of variation across individuals of about 0.20) and considerable uncertainty as reflected in the interquartile range. Americans' individual uncertainty is larger than actual interquartile ranges, whereas for Swiss students individual uncertainty is smaller than actual dispersion. The Swiss students are asked for individual wage expectations under two scenarios, secondary education and tertiary education. These estimates are not significantly different from actually observed medians for these groups (the mean signed error is not significantly different from zero). Most interestingly, the deviations do not differ between respondents with different actual schooling choice (students in high school in a business college or a university of applied science), neither for men nor for men, neither for age 30 nor for age 40. This does not point at selectivity related expectations. Rather, the outcomes are compatible with all students anchoring their expectations on actually observed wages.

Hartog and Webbink (2004) report that first-year students can predict mean salaries by education but they cannot predict their own starting salary after graduation, four years ahead: the correlation between an individual's prediction and the realisation is 0.06.

Indirect indicators, conditional on econometric modelling, are usually somewhat more optimistic but still ascribe a large share of variance to risk. Arcidianoco (2004) estimates that 50-60% of the variance of an ability indicator related to major in university is noise. Cunha, Heckman and Navarro (2005) claim that 60% of variability in returns to education is forecastable at the individual level, and, hence, related to individual heterogeneity, leaving 40% for risk. Perhaps it is not too bold at this stage to say that direct evidence supports the assumption that individual risk may well be approximated by observed (residual) variance while econometric modelling lends more support to a selectivity problem.

The more conventional, two-stage approach to selectivity correction is restricted to a few alternative educations, as these models are not easily estimated for, say, 30 or more alternatives. Chen and Khan (2005) use the 1976 observations from the NLSM 1966 data to assess residual variance with just two options, high school and with college education. Straightforward OLS estimation of an earnings model yields residual standard deviation of 0.397 for college and 0.370 for high school graduates, while two-stage Heckman estimation gives 0.455 and 0.445 (o.c., Table 5). The effect of the selectivity correction on the relative dispersion is quite modest: one is underestimated by 13%, the other by

17% and the college to high school variance ratio falls from 1.07 to 1.02 after correcting for selectivity. More elaborate semi-parametric estimation methods based on matching propensities, gives somewhat larger but still not dramatic effects.

Chen (2006, Table 4) uses the NLSY 1979 panel to estimate heterogeneity and uncertainty in wage variance for four categories of education: less than high school., high school, some college, college graduate. Potential wage inequality is calculated as the transitory component in the residual variance plus the permanent component corrected for selectivity and truncation. It surpasses observed variance by 2 to 13% (13, 7, 2 and 12%, respectively). She defines wage uncertainty (or risk, as I would call it[12]) as the variance of the transitory component and of the permanent component after deducting the effect of reward for unobserved school factors (the correlation between wage residual and schooling choice residual). Uncertainty (the variance of the wage component that the individual cannot foresee when choosing an education) is 1 to 8 % lower than observed residual variance (the ratio's of uncertainty to observed variance are 0.95, 0.99, 0.99 and 0.92, respectively).

We conclude that direct measurement of potential students' information points to a large dispersion between individuals and substantial perceived uncertainty by individuals on their own prospects. The evidence is not obviously inconsistent with students using crude observations on cohorts presently active in the labour market to predict the future effects of their choices. Structural econometric modeling suggests that individuals have more

---

[12] According to Cochrane (2001) this is improper if consumption smoothing is feasible.

precise information on the pay-offs to their own actions, but in several instances the correction for selectivity on measures of risk as used here is negligible.

While one may just claim that the need for selectivity correction is not evident, a negative test result would no doubt be more convincing to its proponents. However, estimating residual earnings variance and skew after correcting for self-selection is not an easy route: it would require estimating selection of an education from at least 30 or 40 alternatives. It would of course also be interesting to study the effect of a truly exogenous shift in expected financial risk associated with education on wages. However, such a case could not yet be identified.

8. Alternative explanations

One might think that the observed relationship reflects something other than risk compensation. One argument might be that earnings distributions obey the lognormal distribution, in which a relationship between mean and variance is inevitable. It is well known that when a variable W has a lognormal distribution with parameters μ and s^2, there is a linear relationship between the mean and standard deviation of W: $E[W] = \exp\left(m + 0.5s^2\right)$ and one may write the variance of W as $Var(W) = \left(E[W]\right)^2 \exp\left(s^2 - 1\right)$. Thus mean and variance of W will both increase if the log variance increases, and even the relative variance (the coefficient of variation will increase). Thus, one might argue that we only have reproduced a property of the lognormal distribution. There are several arguments to counter this interpretation. First, lognormality does not explain anything: even if we were to observe that a variable perfectly obeys a lognormal distribution, we still want to know why this is so. Risk compensation might be precisely the argument. Second, lognormality is not an iron law and statistical tests often reject it (as we show in Hartog and Vijverberg, 2003 for our type of data). Third, in Diaz-Serrano, Hartog and Nielsen (2003) we find risk compensation in earnings also if we measure risk as the intertemporal variation in earnings from panel data; in this case there is even less reason to believe that we just reproduce an iron law of distribution. Fourth, a reasonable model of schooling choices and resulting earnings distributions by education does not imply a necessarily positive relationship between mean and variance. Finally, the most compelling counter-argument is that it would require skew also to be positively related to the mean, as skew for a lognormal variable equals $(2+\exp s^{2})\sqrt{\exp s^{2}-1}$. However, we consistently find a negative relationship.

The underlying argument might also be a price effect on real skills. If earnings emerge as the reward for a skill, with the price of the skill determined in the labour market, an increase in the market price of the skill, with stable underlying distribution, will increase both mean and variance of earnings. But if so, the third moment of the earnings distribution should also go up.

Tournament theory (Lazear and Rosen, 1981) also implies a relationship between mean and variance: if shocks to performance become more important, both mean and variance of wages will be affected. With risk neutral agents, optimal effort is independent of the variance of the shocks to output[13], while the dispersion between the two incomes increases if the shock variance increases[14]. Hence, as one would expect, if agents are risk neutral, there is no compensation for risk. In case of risk averse agents, the relationship between mean and dispersion may be positive or negative, depending on parameter values (see Appendix 2).

Job search theory has a systematic explanation for the persistence of wage differences among observationally identical individuals, pointing to the cost of mobility and of searching for information. A distribution with a large dispersion presents ample opportunity to find a high wage, and one might think the theory yield the same predictions as the risk compensation model. This is not the case however. In the simplest possible model, unemployed workers sample from a wage offer distribution with a reservation wage. As Mortensen (1986, p 864) shows, an increase in the dispersion of the wage offer distribution increases the reservation wage. If the wage offer distribution is normal, an increase in the truncation level will increase the mean of the observed wage distribution (realised offers) while the effect on the variance can not be unambiguously signed.[15] Thus, this model does not unambiguously predict the observed positive relation ship between mean and variance of observed wages. In a more informal sense one might point to the appreciation that individuals will have for a higher dispersion in the job offer distribution, as this will bring them higher expected wage. This would have a negative

[13] Equation (9), Lazear and Rosen (198l) and page 31 in Lazear (1995).
[14] Equation (10), Lazear and Rosen (1981).
[15] See Maddala (1983), p.365. The conditional mean M of the truncated normal distribution is positive in threshold c. The conditional variance V= 1 – M(M-c) cannot be signed as M is positive in c but M-c is negative, as the derivative of M to c is smaller than 1.

effect on mean wages, as supply increases. If the higher wages is to be realised from repeated search, after entering the labour market, and if the higher variance in the offer distribution translates into higher variance in the observed wage distribution, one may perhaps expect a lower starting wage in high variance educations, as individuals appreciate the opportunity to search in a distribution with great opportunities, and a higher mean wage after sufficient experience, when search has paid off. Thus, while in this case predictions are identical for advanced careers, the effect on early careers stages has the power to discriminate. The SEO/Elsevier data analysed in Berkhout, Hartog and Webbink (2006) are restricted to starting wages, 1 to 2 years after graduating from tertiary education. As Table 3 above shows, starting wages respond positively to risk (variance), which does not support the merger of risk preferences and search theory. The role of skew in search theory has not been investigated. Intuitively, one expects opposing effects: the opportunity to arrive in the high end of the distribution, by prolonged search, will push up mean earnings. But as individuals like this opportunity, they will accept a lower wage. Again, one might differentiate by career stages. At young ages, the effect of increased supply dominates and the wage will be depressed. Later, when individuals have moved up through search, the wage should be boosted. But as pointed out above, the estimated effects for starters in the labour market have no different signs[16].

[16] One commentator has remarked that workers can continue searching on the job and that there is repeated sampling rather than a single draw. This would imply that risk is not properly measured by observed wage variance. However, the argument is more complex. Repeated sampling keeps chopping off the lower end of the offer distribution, with presently earned wages as threshold. If the variance of the offer distribution increases, this will increase the mean of observed wages. But as the effect on observed variance cannot even be predicted for a single draw, the effect on observed variance in this case will be even harder to predict. Hence, we cannot predict the relationship between mean and variance of observed wages.

9. What have we learned?

Our empirical results have quite convincingly shown that wages in an occupation/education relate positively to the variance and negatively to the skew within that occupation/education. These results are compatible with risk averse individuals demanding compensation for risk and willing to pay for a chance to obtain really high wages. There are many supporting arguments that make this a plausible link. Omitted ability is not a problem. The best test to check whether the RAM relation is a truly causal relationship would be to analyse the consequence of a convincing exogenous shift in risk. Unfortunately, we have not yet been able to identify such a shift.

The magnitudes of the elasticities are not large. Interestingly, including the risk compensation terms in the Mincer equation has generally no effect on the estimated rate of return to education. This indicates that the estimated return does not suffer from omitted variable bias if risk is ignored, counter to the suggestions of Weiss (1972), Olson, White and Shefrin (1979) and Low and Ormiston (1991). The outcome is compatible with the finding that (residual) earnings dispersion within schooling categories has no robust standard pattern in relation to the level of schooling. Apparently, compensation for postponing earnings and compensation for risk are uncorrelated.

References

Berkhout, P. and J. Hartog (2006), Starting wages respond to employer risk, Working Paper Universiteit van Amsterdam

Berkhout, P., J. Hartog and D. Webbink (2006), Risk compensation under worker heterogeneity, Working Paper Universiteit van Amsterdam

Blanchard, O. J. and S. Fisher (1989), *Lectures on macroeconomics*, Cambridge, Mass: MIT Press

Brunello, G., C. Lucifora and R. Winter-Ebmer (2004), The wage expectations of European business and economics students, *Journal of Human Resources*, 39 (4), 1116-1142

Card. D. (1999), The causal effect of education on earnings, in O. Ashenfelter and D. Card (editors), *Handbook of Labor Economics*, Chapter 30, volume 3A, Amsterdam: North-Holland

Chen, S. (2006), Estimating the variance of wages in the presence of selection and unobserved heterogeneity, Working Paper SUNY-Albany

Chen, S. and Khan (2005), Estimating the causal effect of education on wage inequality using IV methods and sample selection models, Working Paper SUNY-Albany

Cochrane, J. (2001), *Asset pricing*, Princeton: Princeton university Press

Davis, S. and P. Willen (2000), "Occupation-level income shocks and asset returns: their covariance and implications for portfolio choice", Working Paper 7905, NBER.

Diaz Serrano, L. (2000), Human capital, progressive taxation and risk aversion, PhD Dissertation, Universitat Rovira i Virgil, Spain

Diaz Serrano, L. (2005), On the negative relationship between labor income uncertainty and homeownership: risk aversion versus credit constraints, *Journal of Housing*, 14 (2), 109-126

Diaz Serrano, L. and J. Hartog (2006), Is there a risk-return trade-off across educations? Evidence from Spain, *Investigaciones Economicas* 30 (2), 353-380

Diaz Serrano, L., J. Hartog and H. Skyt Nielsen (2004), Compensating wage differentials for schooling risk in Denmark, Discussion Paper Maynooth/Amsterdam/Aarhus, an earlier version was published as IZA DP 963 in 2003; revision 2007

Dohmen, T., A. Falk, D. Huffman en U. Sunde (2006), Individual risk attitudes: new evidence from a large representative experimentally validated survey. IZA Discussion Paper 1730

Feinberg, R.M. (1981), Earnings-risk as a compensating differential, *Southern Economic Journa*l, 48, 156-163.

Garrett, T.A. and R.S. Sobel (1999), Gamblers favor skewness, not risk: further evidence from United States' Lottery games, *Economics Letters*, 63, 85-90.

Golec, J. and M. Tamarkin (1998), Bettors love skewness, not risk, at the horse track, *Journal of Political Economy*, 106 (1), 205-225.

Gollier, C. (2001), *The economics of risk and time*, Cambridge, Mass: The MIT Press

Hartog, J., E. Plug, L. Diaz Serrano and J. Vieira (2003), Risk compensation in wages, a replication, *Empirical Economics*, 28, 639-647

Hartog, J. and W. Vijverberg (2006), On compensation for risk aversion and skewness affection in wages, *Labour Economics* Special Issue on Education and Risk, forthcoming

Hartog, J. and W. Vijverberg (2007), Schools, skills and risk, *Economics of Education Review*, forthcoming

Jacobs, B., J. Hartog and W. Vijverberg (2005), Self-selection bias in estimated wage premiums for earnings risk, Working Paper Universiteit van Amsterdam, under revision

Johnson, W. (1977), Uncertainty and the distribution of earnings, in: F. Juster (ed.), *The distribution of economic well-being*, Cambridge, Mass: Ballinger for NBER, 379-396

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

Kodde, D. (1985), Microeconomic analysis of demand for education, PhD thesis Erasmus Universiy Rotterdam

Lazear (1995), *Personnel Economics,* Cambridge, Mass: MIT Press

Lazear, E (2005), Entrepreneurship, *Journal of Labor Economics*, 23 (4), 649- 680

Lazear, E. and S. Rosen (1981), Rank-order tournaments as optimal labor contracts, Journal of Political Economy, 89 (5), 841-864

Levhari, D. and Y. Weiss (1974), The effect of risk on the investment in human capital. *American Economic Review*, 64 (6), 950-963.

Low, S. and M. Ormiston (1991), Stochastic earnings functions and risk, *Southern Economic Journal*, 57 (4), 1124-1132

Ma, X, (2005), Research on the risk of individual investment in education--using the Urban Chinese in the 1990s as an example, PhD dissertation Peking University

Maddala, G. s. (1983), *Limited-dependent and qualitative variables in econometrics*, Cambridge: Cambridge University Press

McGoldrick, K. (1995), Do women receive compensating wages for earnings risk? *Southern Economic Journal*, 62, 210-222.

McGoldrick, K. and J. Robst (1996), The effect of worker mobility on compensating wages for earnings risk, *Applied Economics*, 28, 221-232.

Mincer, J. (1974; reprinted 1993), *Schooling, experience and earnings*, , Aldershot UK: Gregg Revivals

Moore, M.J. (1995), Unions, employment risks, and market provision of employment risk differentials, *Journal of Risk and Uncertainty*, 10(1), 57-70.

Mortensen, D. (1986), Job search and labor market analysis, Chapter 15 in O. Ashenfelter and R.Layard (eds), *Handbook of Labor Economics*, Volume 2, Amsterdam: North-Holland

Moulton, B. R. (1986), Random group effects and the precision of regression estimates, *Journal of Econometrics*, 32, 385-397

Murphy, K. and R. Topel (1985), Estimation and inference in two-step econometric models, *Journal of Business and Economic Statistics*, 3, 88-97

L.Olson, H. White and H. Shefrin (1979), Optimal Investment in Schooling When Incomes are Risky, *Journal of Political Economy*, 87 (3), 522-539

Palacios-Huerta, I.(2003), An empirical analysis of the risk properties of human capital returns, *American Economic Review* 93 (3), 948-964

Pereira, P and P. Martins (2002), Is there a return-risk link in education? *Economics Letters*, 75 (1), 2002, 31-37

Raita, S. (2005), The risk of investment in human capital, PhD thesis Tinbergen Institute Universiteit van Amsterdam

Rosen, S. (1986), The theory of equalizing differences, in O. Ashenfelter and R. Layard (editors), *Handbook of Labor Economics*, Chapter 12, volume 1, Amsterdam: North-Holland

Saks, R. and S. Shore (2003), Risk and career choice, paper presented at the AEA Meetings, San Diego, January 2004

Shaw, K.L. (1996), An empirical analysis of risk aversion and income growth, *Journal of Labor Economics* 14 (4), 626-653.

Weiss, Y. (1972), The risk element in occupational and educational choices, *Journal of Political Economy*, 80 (6), 1203-1213

Wolter, S and B. Weber (2004), Returns to education: are students' expectations rational? Working Paper Swiss Coordination Center for Research in Education, Aarau, Switzerland

Appendix 1. Risk compensation in Levhari and Weiss (1974)

Equation (5) in Levhari and Weiss (1974), in their notation, reads

$$E\left\{\frac{\partial U}{\partial c_1}\left[\,f_l-(1+r)y_0\,\right]\right\}=0$$

or

$$\frac{E\dfrac{\partial U}{\partial c_1}f_l}{y_0\,E\dfrac{\partial U}{\partial c_1}}=1+r$$

subtracting $E(f_l)/y_o$ on both sides, and using $E(xy)=E(x)E(y)+\mathrm{cov}(x,y)$, we can write

$$\frac{E\,f_l}{y_0}=1+r-\frac{\mathrm{cov}\left(\dfrac{\partial U}{\partial c_1},f_l\right)}{y_0\,E\dfrac{\partial U}{\partial c_1}}$$

Kodde (1985, Chapter 7) shows, using a second-order Taylor expansion around $E(m)$ that we may write

$$\mathrm{cov}\left(\frac{\partial U}{\partial c_1},f_l\right)=\frac{\partial^2 U}{\partial c_1^2}\mathrm{f}_m\,\mathrm{f}_{lm}\,s^2-\left|\frac{\partial^3 U}{\partial c_2^3}\mathrm{f}_m^2\,\mathrm{f}_{lmm}+\frac{\partial^2 U}{\partial c_2^2}\,f_{mm}\,f_{lmm}\right|s^4/4$$

where s^2 is the variance of m. Assuming $f_{lmm}=0$, as Levhari and Weiss (1974) implicitly do, we can write

$$E\frac{f_l}{y_0}=1+r-y_0\frac{\partial^2 U/\partial c_1^2}{E\partial U/\partial c_1}\mathrm{f}_m\,\mathrm{f}_{lm}\,\frac{s^2}{y_0^2}$$

$$=1+r+r\,\mathrm{f}_m\,\mathrm{f}_{lm}\,\frac{s^2}{y_0^2}$$

where r is relative risk aversion evaluated at expected values. Thus, with risk aversion positive, and f_{lm} positive as assumed by Levhari and Weiss, under uncertainty the expected return on human capital surpasses the Mincer rate by a term that is proportional to risk aversion and relative variance.

Appendix 2. Mean and dispersion in the tournament model (Lazear and Rosen, 1981)

Under risk aversion, equilibrium condition (24) in Lazear and Rosen (1981) can be written as

$$m = \frac{1}{C'(m)} \frac{V}{1 + as^2} \text{ with } a = sC''\Pi > 0$$

Totally differentiating and rewriting yields

$$\frac{dm}{ds} = -\frac{2as\, m/\left(1 + as^2\right)}{1 + mC''(m)/C'(m)} < 0$$

This holds because C’ > 0 and C” > 0 by assumption. One might also write the equilibrium condition as

$$mC'(m) = \frac{V}{1 + as^2}$$

With the LHS increasing in μ, an increase in s^2 has to be matched by a decrease in μ.

The equilibrium wage dispersion, according to footnote (9) in Lazear and Rosen (1981) obeys

$$C'(m) = wbs^{-1} \text{ with } w = w_1 - w_2 \text{ and } b = 1/\sqrt{2\Pi}$$

Totally differentiating this condition (to μ, w and s) yields,

$$\begin{aligned} \frac{dw}{ds} &= \frac{s}{b}\left\{C''(m)\frac{dm}{ds} + wbs^{-2}\right\} \\ &= \frac{mC''(m)}{b}\left\{e_s^m + \frac{1}{mC''(m)/C'(m)}\right\} \end{aligned}$$

where e_s^m is the elasticity of μ to s. Thus

$$\frac{dw}{ds} > 0 \text{ if } e_s^m > -\left[\frac{C''(m)}{C'(m)} m\right]^{-1}$$

i.e. if the elasticity of effort is between zero and the inverse of what might be called relative effort aversion. The elasticity is endogenous, characterizing the dislocation of equilibrium effort. If we substitute from equation (1) above, we can derive

$$\frac{dw}{ds}=\frac{C'(m)}{b}\left\{1-\frac{2as^2 mC'(m)}{1+mC''(m)/C'(m)}\frac{1}{V}\right\}$$

This is ambiguous in sign, depending on exogenous parameter values.